

Seq primer: TF
Class: Shotgun.
Location/Qualifiers

1. .693
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="LANDSBERG ERECTA"
/db_xref="taxon:3702"
/clone="LBERM21"
/clone_lib="LBERG"
/note="Organ: Leaf; Vector: pUC19K; Total genomic DNA was sheared to 0.6-0.8 Kbp before ligation."

ORIGIN

Query Match 18.8%; Score 614.4; DB 28; Length 693;
Best Local Similarity 99.8%; Pred. No. 1.3e-99;
Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
1 (bases 1 to 613)
Bell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utechtach, T., Feldbylun, T., Liang, F., Creasy, T., and Frazer, C.M.
Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
Unpublished (2000)
Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: ac@tigr.org
For additional information, see <http://www.tigr.org/db/at.html>
Seq primer: TF
Class: Shotgun.

FEATURES

location/Qualifiers
1. .613
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="LANDSBERG ERECTA"
/db_xref="taxon:3702"
/clone="LBERG20"
/clone_lib="LBERG"
/note="Organ: Leaf; Vector: pUC19K; Total genomic DNA was sheared to 0.4-0.7 Kbp before ligation."

ORIGIN

Query Match 18.6%; Score 608.2; DB 28; Length 613;
Best Local Similarity 99.5%; Pred. No. 1.7e-98;
Matches 610; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 829 AGTCATTTCCAGAGAGAAACAACCAACGAGAAACACTTCTGAAAAAATTGGGTTT 888
DB 688 AGTCATTTCCAGAGAGAAACAACCAACGAGAAACACTTCTGAAAAAATTGGGTTT 629
QY 889 CTCAGATTCGAATCTCTTGAAGCTTGGGGTGTCTACGACCAAGGAGGTACCAAG 948
DB 628 CTCAGAAATTCGAATCTCTTAAAGCTTGGGGTGTCTACGACCAAGGAGGTACCAAG 569
QY 949 TATCTGGGAGAGAAACAATGTATCCGGAGAGGTAAACGCAATGTCTGGGGCAATTCAAG 1008
DB 568 TATCTGGGAGAGAAACAATGTATCCGGAGAGGTAAACGCAATGTCTGGGGCAATTCAAG 509
QY 1009 CTAAATATCTGCTGGGAGAGAGCTTGTAGCAGAAATATATTAACAATTGTGGCAC 1068
DB 508 CTAAATATCTGCTGGGAGAGAGCTTGTAGCAGAAATATATTAACAATTGTGGCAC 449
QY 1069 CCCCACTGTATCTTCTGAGAGCTTCTTGGAGAGAGATGGAATTGGAGGAGAGGTT 1128
DB 448 CCCCACTGTATCTTCTGAGAGCTTCTTGGAGAGAGATGGAATTGGAGGAGAGGTT 389
QY 1129 CTGCTCAGACACAGCTGTGAGAGAGTTCTGACGTGAGAGATGATGTATATGCTT 1188
DB 388 CTGCTCAGACACAGCTGTGAGAGAGTTCTGACGTGAGAGATGATGTATATGCTT 329
QY 1189 CTGAGAGAGAGATGATCCGATGCTTGTGATGATGATGATGATGATGATGATGATG 1248
DB 328 CTGAGAGAGAGATGATCCGATGCTTGTGATGATGATGATGATGATGATGATGATG 269
QY 1249 ATTATGATCGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1308
DB 268 ATTATGATCGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 209
QY 1309 AAAAGTTCTTGGAGCTTGGATAGCTTGTGATGATGATGATGATGATGATGATGATG 1368
DB 208 AAAAGTTCTTGGAGCTTGGATAGCTTGTGATGATGATGATGATGATGATGATGATG 149
QY 1369 AGTGGCATTTGTCAGCTTGTGAGAAACGAGACCTGTGTCATGATGATGATGATGATG 1428
DB 148 AGTGGCATTTGTCAGCTTGTGAGAAACGAGACCTGTGTCATGATGATGATGATGATG 89
QY 1429 CTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1444
DB 88 CTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 73

RESULT 2
AO964580/c 613 bp DNA linear GSS 28-JAN-2000
LOCUS LERGX20TF LERGX Arabidopsis thaliana genomic clone LERGX20, genomic

DEFINITION
Survey sequence.
ACCESSION
AO964580
VERSION
AO964580.1 GI:6792281
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana

QY 608 GTTGGCTTGTATTAATCTCAAGTTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 667
DB 613 GTTGGCATGTTATTAATCTCAAGTTCATGTCATGTCATGTCATGTCATGTCATGTCATG 554
QY 668 GTTCTTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 727
DB 553 GTTCTTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 494
QY 728 AAAGACGTCAGAGGTGATATAGGCTGAGGTGAAACATGTTGTTCAAGTTTGGCAGG 787
DB 493 AAAGACGTCAGAGGTGATATAGGCTGAGGTGAAACATGTTGTTCAAGTTTGGCAGG 434
QY 788 GACGAGCTGCTTCTTCAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 847
DB 433 GACGAGCTGCTTCTTCAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 374
QY 848 CAAGAAACAAACGAGAAACCTTCTGAAACAACTTGGGTTCTCAAGATTCGAATCTCTCC 907
DB 373 CAAGAAACAAACGAGAAACCTTCTGAAACAACTTGGGTTCTCAAGATTCGAATCTCTCC 314
QY 908 TAGAGCTTGGGGTGTGTCAG 967
DB 313 TAGAGCTTGGGGTGTGTCAG 254
QY 968 TGTATCCGGAGAGATGACGCAATGTCGCGGCAATTCAGACTAATATCTGTCGGGG 1027
DB 253 TGTATCCGGAGAGATGACGCAATGTCGCGGCAATTCAGACTAATATCTGTCGGGG 194
QY 1028 ACGAGGCTTGAAGAGAAAGTATATACAACTTGTGGACCCCCCACTGATCTGCCCC 1087
DB 193 ACGAGGCTTGAAGAGAAAGTATATACAACTTGTGGACCCCCCACTGATCTGCCCC 134
QY 1088 TCCCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1147
DB 133 TCCCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 74
QY 1148 GCAAGAGTTTCTGACGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1207

Db 73 GCAGAGATTTCTGACGTGAGAGATGATGTGTAATGCTTCTGAGAGAAAGAAATGATTC 14
 Qy 1208 CGATGCTTGGAT 1220
 |||||
 Db 13 CGATGCTTGGAT 1

RESULT 3
 AU236368
 LOCUS
 DEFINITION AU236368 RAF14 Arabidopsis thaliana cDNA clone RAF14-93-K05 5', mRNA sequence.
 ACCESSION AU236368
 VERSION AU236368.1 GI:19875537
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 647)
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y., and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 Unpublished (2002)
 TITLE Contact: Motoaki Seki
 JOURNAL Plant Functional Genomics Research Group
 COMMENT RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rtc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda PhiC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified plasmid vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES
 source
 1. 647
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF14-93-K05"
 /issue_type="root"
 /lab_host="DH10B"
 /clone_1lb="RAF14"
 /note="Site_1: BamHI, Site_2: SalI"

ORIGIN
 Query Match 16.5%; Score 542; DB 9; Length 647;
 Best Local Similarity 99.5%; Pred. No. 1.1e-86;
 Matches 564; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 687 GTTGTAAAAATGAGTTCTAGAGCTGCTCAATGCTTAAGAAAAAGACGTTCAAGGCTGCT 746
 Db 83 GTTGTAAAAATGAGTTCTAGAGCTGCTCAATGCTTAAGAAAAAGACGTTCAAGGCTGCT 142

Qy 747 TATAGGCTGAGGTTGAACAGTTGTTCAAGTTTGGCAGGAGCAGAGCTGCTTCTTCA 806
 Db 143 TATAGGCTGAGGTTGAACAGTTGTTCAAGTTTGGCAGGAGCAGAGCTGCTTCTTCA 202

Qy 807 CAAGATGATGAGAGAGAGTGGAGGTCATTTCCAAAGAAACAAGAACCAACAGAGAAC 866
 Db 203 CAAGATGATGAGAGAGAGTGGAGGTCATTTCCAAAGAAACAAGAACCAACAGAGAAC 262

Qy 867 ACTTCTGAAAAAATCTGGGTTTTCAGAAATTCGAATCTCTCTAGAGCTTGGGGTGGTCAG 926
 Db 263 ACTTCTGAAAAAATCTGGGTTTTCAGAAATTCGAATCTCTCTAGAGCTTGGGGTGGTCAG 322

Qy 927 CAGCAAGGAGAGGTAGCAACGTATCTGGAGAGAAAAACAATGATCCGGAGAGTAAAC 986
 Db 323 CAGCAAGGAGAGGTAGCAACGTATCTGGAGAGAAAAACAATGATCCGGAGAGTAAAC 382

Qy 987 GGCATGCTGGGGCATTCAAGCTAACATATCTGGTGGGGAGCAGAGCTTGAAGAAAG 1046
 Db 383 GGCATGCTGGGGCATTCAAGCTAACATATCTGGTGGGGAGCAGAGCTTGAAGAAAG 442

Qy 1047 TATGATACAACTTTGGGACACCCCACTGATCTCGCCCTCTTTGGAAGAGAGATGG 1106
 Db 443 TATGATACAACTTTGGGACACCCCACTGATCTCGCCCTCTTTGGAAGAGAGATGG 502

Qy 1107 AATTGGCAGGCAAGAGAGAGTTCTGTCAGACACACAGCTGTGACAGATTCTTGACGTG 1166
 Db 503 AATTGGCAGGCAAGAGAGAGTTCTGTCAGACACACAGCTGTGACAGATTCTTGACGTG 562

Qy 1167 GAGGATGATGTGATTAATGCTTCTGAGAAAGAGATGATTCGATGCTTTGGATGATCT 1226
 Db 563 GAGGATGATGTGATTAATGCTTCTGAGAAAGAGATGATTCGATGCTTTGGATGATG-TCT 620

Qy 1227 GATGAGACCTTGCAAGTATGATTAAT 1253
 Db 621 GATGAGACCTTGCAAGTATGATTAAT 647

RESULT 4
 AQ959657/c
 LOCUS LEREM21F LERE Arabidopsis thaliana genomic clone LEREM21, genomic survey sequence.
 DEFINITION AQ959657
 ACCESSION AQ959657
 VERSION AQ959657.1 GI:6787358
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 606)
 Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Uterbach, T., Feldblum, T., Jiang, F., Creasy, T., and Frazer, C.M.
 Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
 Unpublished (2000)
 TITLE Contact: Xiaoying Lin
 JOURNAL The Institute for Genomic Research
 COMMENT 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: at@tigr.org
 For additional information, see http://www.tigr.org/cdb/at.ac.html
 Seq primer: TF
 Class: shotgun.

FEATURES
 source
 1. 606
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="LANDSBERG ERRECTA"
 /db_xref="taxon:3702"
 /clone="LEREM21"
 /clone_1lb="LERE"
 /note="Organ: Leaf; Vector: pUC19JK; Total genomic DNA was sheared to 0.6-0.8 kbp before ligation."

ORIGIN
 Query Match 16.3%; Score 534; DB 28; Length 606;
 Best Local Similarity 100.0%; Pred. No. 3e-85;
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 911 AGCTTGGGTGTGTGACGACGAGGAGAGGTAGCAACGTATCTGGAGAGAAAAACAATGT 970
 Db 606 AGCTTGGGTGTGTGACGACGAGGAGAGGTAGCAACGTATCTGGAGAGAAAAACAATGT 547

ORIGIN	Query Match	15.6%;	Score 511.8;	DB 28;	Length 523;
	Best Local Similarity	98.7%;	Pred. No. 2.9e-81;		
	Matches 516;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0
QY	838	CCAAGAAGAACAGAACAAACACAGAGAAACACTTCTGAAAAAATTGGGTTTCTCAGATT	897		
Db	1	CCAAAGAAAGAACAGAACAAACCATGAAACACTTCTGAAAAAATTGGGTTTCTCAGATT	60		
QY	898	CGAATCCCTCAGAGCTTGGGGTGTCAGACGACGAAGGAGAGGTAGCAACGATATGGGA	957		
Db	61	CGAATCCTCTAGAGCTTGGGGTGTCAGCAACAGGAGAGGTAGCAACGATATGGGA	120		
QY	958	GAGGAACCAATGTATCCGGGAGAGGTAAACGGCAATGCTGGGGCAATTCMACTTAACATAT	1017		
Db	121	GAGGAACCAATGTATCCGGGAGAGGTAAACGGCAATGCTGGGGCAATTCMACTTAACATAT	180		
QY	1018	CTGGTCGGGGGACGACCGTTGACGAAAGTATGATAAACAATTGTGGCACCCCCACCTG	1077		
Db	181	CTGGTCGGGGGACGACCGTTGACGAAAGTATGATAAACAATTGTGGCACCCCCACCTG	240		
QY	1078	TATCTGSCCTCTCTTTGAAAGAGAGATGAAATTGGACGACAAAGAGAGTTCTGCTCAGC	1137		
Db	241	TATCTCCCTCTCTTTGAAAGAGAGATGAAATTGGACGACAAAGAGAGTTCTGCTCAGC	300		
QY	1138	ACAACAGCTGACGAGATTTCTGACGTTGAGAGATATGTTGATTAATGCTTCTGAGGAG	1197		
Db	301	ACAACAGATGTCAGAGATTTCTGACGTTGAGAGATATGTTGATTAATGCTTCTGAGGAG	360		
QY	1198	AGAAATGATCCGATGCTTTGGATGATTTCTGATGCGACCTTGCAGATGATGATTATGACT	1257		

RESULT 6	CNS00P9H	457 bp	DNA	Linear	GSS: 28-JUN-1998
LOCUS	CNS00P9H				
DEFINITION	Arabidopsis thaliana genome survey sequence T7 end of BAC F8G21 of TGF library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.				
ACCESSION	AL084227				
VERSION	AL084227.1	GI:5285367			
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 457)				
AUTHORS	Salanoubat, M., Choisme, N., Artiguenave, F., Brottier, P., Wincker, P., Smeaton, D., Saurin, W., Weissenbach, J., and Quetier, F.				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 457)				
AUTHORS	Genoscope.				

```

TITLE      Direct Submission
JOURNAL    Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sege@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
FEATURES   Location/Qualifiers
            1..457
             /organism="Arabidopsis thaliana"
             /mol_type="genomic DNA"
             /strain="Columbia"
             /db_xref="taxon:3702"
source

```



```

/clone="F8G21"
/clone_lib="IGF"
/note="end : T7"

```

Query Match	13.8%	Score 451.2;	DB 29;	Length 457;
Best Local Similarity	99.1%;	Pred. No. 2e-70;		
Matches 453;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

OY	1676	TTTACCTCTTAATTTTTTCCTTGATTCTCACTGATCTTAGAAGTGTACATTTGATGGC	1735
Db	1	TTACTNCTTAAATTTTTCTCTTGATTTCTACTGATCTTAGAAGTGTACATTTGATGGC	60
OY	1736	TCCGCATBGCACAACMAGAGCTGCTGGAATTCCTGCACAAGTATBAGGCTCTTAGAGCAC	1795
Db	61	TCCGCATBGCACAACMAGAGCTGCTGGAATTCCTGCACAAGTATBAGGCTCTTAGAGCAC	120
OY	1796	GCCATTTCCATAAGTGTCCAAGGGGCCATCGTGGGATACAGTGTTCTGATGTTTGAGACAGT	1855
Db	121	GCCATTTCCATAAGTGTCCAAGGGGCCATCGTGGGATACAGTGTCTGATGTTTGAGACAGT	180
OY	1856	CCACTGCGTATTTTGAGAGCCGCAACGCTCCACCCGGAGTTAGCTGAGATGGGGTTAGATA	1915
Db	181	CCACTGCGTATTTTGAGAGCCGCAACGCTCCACCCGGAGTTAGCTGAGATGGGGTTAGATA	240
OY	1916	GAATTTGCTGGGGGTCAAGAAGCGACGATATGTTTTCTGAGAGTGTTCGCCACTGTATGGCT	1975
Db	241	GAATTTGCTGGGGGTCAAGAAGCGACGATATGTTTTCTGAGAGTGTTCGCCACTGTATGGCT	300
OY	1976	TCCTTGCACAAGACGAGATCTGGACATATTCATCAACCTCTCAAAGTTCCTCCCCC	2035
Db	301	TCCTTGCACAAGACGAGATCTGGACATATTCATCAACCTCTCAAAGTTCCTCCCCC	360
OY	2036	AAAAGAAATTTGATATATGCTTTTAGTGTTCATTTGCAATGGAATTTAAAGTTTTGTGTCCGT	2095
Db	361	AAAAGAAATTTGATATATGCTTTTAGTGTTCATTTGCAATGGAATTTAAAGTTTTGTGTCCGT	420
OY	2096	GTTAATGCACTGTTATGTATATATCATATGATTCATT	2132
Db	421	GTTAATGCACTGTTATGTATATATCATATGATTCATT	457

RESULT 7	BH535855	697 bp	DNA	linear	GSS 14-DEC-2001
LOCUS	BH535855				
DEFINITION	BH535855 BOHK Brassica oleracea genomic clone BOHKR13, genomic survey sequence.				
ACCESSION	BH535855				
VERSION	BH535855.1	GI:17767651			
KEYWORDS	GSS.				
SOURCE	Brassica oleracea				
ORGANISM	Brassica oleracea				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euarysta; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.				
REFERENCE	1 (bases 1 to 697)				
AUTHORS	Town,C.D., Van Aken,S., Uterback,T., Koo,H. and Frazer,C.M.				
TITLE	Whole genome shotgun sequencing of Brassica oleracea				
JOURNAL	Unpublished (2001)				
COMMENT	Other_GSSs: BOHKR13TR				

```

FEATURES
  source      Location/Qualifiers
              1. .697
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"

```

```

/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOHk13"
/clone_lib="BOHk"
/notes="Vector: pHOSt; Site 1: BscXI, 2-3 kb sheared
genomic DNA inserted into pHOSt using BstXI linkers"

```

Query Match	13.7%	Score 449.2;	DB 28;	Length 697;
Best Local Similarity	80.2%;	Pred. No. 3.9e-70;		
Matches 558; Conservative	0;	Mismatches 123;	Indels 15;	Gaps 2;

Oy	1302	TGTTTCAAAAAGTCTTTGGCAGCTTGGATAGCTGTGCATCGACGAGATTAATGAACCA	1361
Db	1	TGTTTAAAGATTTTTTCGACAGCTTGGATAGCTGTGTCTATCGACGAGATTAACGACCA	60
Oy	1362	CAGAGCAGTGCATTTGCTCAGCTTTCAGAACGACCTGGTGCATCGATTGGTATPAC	1421
Db	61	CAGAGGCAGTGGCATTTGTCGGGCTTGCAGAACGGGCTGGCGGTATTGACTGGTACAA	120
Oy	1422	CTGCACCCCTTACTAGCTCATGCGAGGACAAAGAGACTTGGCGGATTTAAGCTTCATAGA	1481
Db	121	CTTCAGCCACTTGTAGCTCACGCGAGGACAAAGGAGACTAGACGTGTTAAGCTCCACAGG	180
Oy	1482	GAATTGGCTGAAGTTTAAAGAAAGACTCAGATGAGAGCGCATCTGTCAATCCCTGT	1541
Db	181	GAACCTGCTAGGTTCTCGACAGGAGCTCAGATGAGGAGAGCATGTGCAATCCCTGC	240
Oy	1542	GGTCAGATTTATGGGCGAGTGGAAAGGCTTGGGTCAGATGAAAAGATTTAAGAAATTGTC	1601
Db	241	GGTAAATTTTACGGGCGAGTGGAAAGGCTTGGGTCAGAGGAAAGACATGATATTGTC	300
Oy	1602	TGCGCTCCAAATGTCATCATATGAATATCTAGACTGAGTAAAGACGATTAAGATPAGGTG	1661
Db	301	TGCGCGCCAAATGGTTCATCATGAATATCTAGGCTGAGCAAGGACGAAAGACAAAGTTC	360
Oy	1662	GAATTCCTTCTGTCTTTACTCTTTAAT-----TTTCTCTGCATCTTACTGAT	1711
Db	361	AACCTTCCTCTCTGTCTTTTCTTTTCTTTTAAATCATCTTGCCTTGCCTTCGAT	420
Oy	1712	-----CTTAGATGTTCATTTGATGAGGCTCCGAGTGGGCAACCAAGAGCTCGAGATA	1766
Db	421	CATACATAAGTTTGTATCATATCATGTGGCTAGAGGATGGCAACCGGAGCTGCTGGAGTA	480
Oy	1767	CTTCGACAAGTATGAGGCTCTTATAGAGCACGCCAATCTTATGTGTCCACAGGGCCATCGTGG	1826
Db	481	CTTTAAAGAGATTCGTCTATTAAAGCACGCCACTCTTATGTGTCCACAGGGCCATCGTGG	540
Oy	1827	GATGAGTGTCTGATGTTTAAAGACAGTGCCTAGCTATTGGAGGGCCGAACGCTTCCA	1886
Db	541	GATTAGTGTCTGATCTTTGAGAGTACCGCCACTAGGCTTATTGAGGCTGACCGCCTTACA	600
Oy	1887	CCGGGAGTTAGCTAGATGGGGTTAAGTAAATTTGCTGGGGGTCAAGAACGCCATATGTT	1946
Db	601	CAGGAGCTTGGCTCAAAAGGGGTTGATTAAGACGCTTGGGACCGTCGCGCAGATGTT	660
Oy	1947	TTCTGAGAGTTCGCGCAACTGATGAGCTTCCTTGC 1982	
Db	661	TTCTGAGAGTTCGCGCAACTGATGAGCTTCCTTGC 696	

RESULT 8	548 bp	RNA	linear	EST 08-SEP-1999
AI999551/c				
LOCUS				
DEFINITION	AI999551			
	701556368 A. thaliana, Columbia Col-0, rosette-3 Arabidopsis			
	thaliana cDNA clone 701556368, mRNA sequence.			
ACCESSION	AI999551			
VERSION	AI999551.1			
KEYWORDS	GI:5846456			
SOURCE	EST.			
ORGANISM	Arabidopsis thaliana (thale cress)			
	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			

rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 548)

REFERENCE
AUTHORS
Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gaxone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kasurty, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L., and Hanson, D.

TITLE
Arabidopsis thaliana Gene Expression Microarray
Unpublished (1999)

COMMENT
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

FEATURES
source
1..548
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultiar="Columbia Col-0"
/db_xref="taxon:3702"
/clone="701556368"
/cisse_type="rosette"
/dev_stage="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, rosette-3"
/note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunted, ligated to SalI adaptor, digested with NotI, size selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN
Query Match 13.3%; Score 434.2; DB 9; Length 548;
Best Local Similarity 84.0%; Pred. No. 2e-67;
Matches 536; Conservative 0; Mismatches 12; Indels 90; Gaps 1;

2488 GCACACGACAGGTTTTCATGATTCATCAAGATCCATGAAGAAGAGAGCAAG 2547
548 GCACACGACAGGTTTTCATGATTCATCAAGATCCATGAAGAAGAGAGCAAG 489
2548 GAGGAGATTTTCGATGTCGACGACGAGAGAGGTCGAGTTTGGCCAGACGAG 2607
488 GAGGAGATTTTCGATGTCGACGACGAGAGAGGTCGAGTTTGGCCAGACGAG 429
2608 CAGAACATTATCCCTTAGCAATGACGATGCCGAAAGAGGTATGTAATACTACA 2667
428 CAGAACATTATCCCTTAGCAATGACGATGCCGAAAGAGGTATGTAATACTACA 390
2668 TAAATCCCTGCGCTTTTGTGTTTCAAACTAAGATTAATTCGCTTTTGA 2727
389 ----- 390
2728 TTCTTTCGAGAGTGAAGAGAGTGAAGTTCATGAGTTTCAAGAAAGAGATGAG 2787
389 -----AGAGCTGAGAGAGAGTGAAGTTCATGAGTTTCAAGAAAGAGATGAG 339
2788 GAGTTTTCGAGAGAGAGAGATGCTGATAAAGATCAAGAGAGAGATGAGACATG 2847
338 GAGTTTTCGAGAGAGAGAGATGCTGATAAAGATCAAGAGAGAGATGAGACATG 279
2848 AAGAGAGAGATCCGAGAGATTTTGTGTCGAGAAAGATTTGATGAGGCTTTGAA 2907
278 AAGAGAGAGATCCGAGAGATTTTGTGTCGAGAAAGATTTGATGAGGCTTTGAA 219
2908 CAGCTCTGTAACAAGATGAGCTTCAATGAAGATGATTGAGACAAATCTGTGAC 2967

DB 218 CAGCTCATGTACAGCATGCGCTTCAATGAAGATGATGAGACAAAGCTGTGAC 159
QY 2368 AAGACAAGACTAGTTCTTTGTTTCTTTGTTGATGTCGAAAGAGATCTGAGA 3027
DB 158 AAGACAAGACTAGTTCTTTGTTTCTTTGTTGATGTCGAAAGAGATCTGAGA 99
QY 3028 GACTCCATTTAATGATGAGACAAATCTAAGAGATTTAGATTATCTCCAAATTT 3087
DB 98 GACTCCATTTAATGATGAGACAAATCTAAGAGATTTAGATTATCTCCAAATTT 39
QY 3088 TAGTAGAGGATCTAAGAGAGATTAAGTCTGTGAC 3125
DB 38 TAGTAGAGGATCTAAGAGAGATTTAATCTGTGAC 1

RESULT 9
BH618393/c
LOCUS
DEFINITION
SALK_039005 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_039005, genomic survey sequence.
ACCESSION
BH618393
VERSION
BH618393.1 GI:18428488
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmermann, J., and Ecker, D.R.
A Sequence-indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)
CONTACT: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated intron of At5g23570.
Class: TDNA tagged.

FEATURES
source
1..443
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_039005"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN
Query Match 12.7%; Score 417.4; DB 28; Length 443;
Best Local Similarity 99.3%; Pred. No. 2.1e-64;
Matches 440; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

1694 CTCTTGACATTTAGATCTTGAATGTTTACATTTGATGCTCGGATGGCAACAG 1753
DB 443 CTCTTGACATTTAGATCTTGAATGTTTACATTTGATGCTCGGATGGCAACAG 384
QY 1754 AGCTGCTGAATACCTTGCAGAGTATGAGCTCTTGAAGACGCAATCTGATGTCAC 1813
DB 383 AGCTGCTGAATACCTTGCAGAGTATGAGCTCTTGAAGACGCAATCTGATGTCAC 324

QY	1814	AGGCGCATCGCGGGANVAGTGTTCCTGATGTTTGAAGAGAGGCCATCGGATTTTGGAGG	18173
Db	323	AGGCGCATCGCGGGANVAGTGTTCCTGATGTTTGAAGAGAGGCCATCGGATTTTGGAGG	264
QY	1874	CCGAAAGCGCTCCA -CCGGAGGTAGTCTGAGATGGGGTTAGATAGAAATTCGCTGGGGTCAG	19322
Db	263	CCGATCGCCTCCANCCGGGAGGTAGTCTGAGATGGGGTTAGATAGAAATTCGCTGGGGTCAG	204
QY	1933	AAGCGCACTAATGTTTCTCGAGAGGTGTTCGCCACTGATAGGCTTCCTTGCAACGAAAGAA	19932
Db	203	AAGCGCACTAATGTTTCTCGAGAGGTGTTCGCCACTGATAGGCTTCCTTGCAACGAAAGAA	144
QY	1993	GATCTGACATATTTCAATCAACACCTCTCA -AGGTCTCTCCGCCAAAGAAATTTGATATTA	2051
Db	143	GATCTGACATATTTCAATCAACACCTCTCAAGGTTCTCTCCGCCAAAGAAATTTGATATTA	84
QY	2052	TGCTTTAGTATTTGTCATTGGAAATTTAAAGTTTGTGTGTCGTCGTTAATGATCATCTGTTA	21111
Db	83	TGCTTTAGTATTTGTCATTGGAAATTTAAAGTTTGTGTGTCGTCGTTAATGATCATCTGTTA	24
QY	2112	TGTATATATCTATGATTCATTNG	2134
Db	23	TGTATATATCTATGATTCATTNG	1
RESULT 10			
AU2272777/c			
LOCUS	AU227277	421 bp	mRNA
DEFINITION	AU227277 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-93-K05 3', mRNA sequence.		
ACCESSION	AU227277		
VERSION	AU227277.1	GI:19741924	
KEYWORDS	EST.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eudicotyledons; Magnoliopsida; Eudicotyledons; core eudicots; Rosids; eustroids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1 (bases 1 to 421)		
AUTHORS	Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.		
	Large scale analysis of Arabidopsis full-length cDNA unpublished (2002)		
TITLE	Contact: Motoaki Seki		
JOURNAL	Plant Functional Genomics Research Group		
COMMENT	RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: msekierc.riken.go.jp		
	An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda FLIC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.		
FEATURES			
source	location/Qualifiers		
	1..421		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/clone="RAFL14-93-K05"		
	/tissue_type="root"		
	/lab_host="DH10B"		
	/clone_1ib="RAFL14"		
	/note="Site_1: BamHI; Site_2: SalI"		
ORIGIN			
Query Match	11.9%;	Score 389;	DB 9; Length 421;

	Best Local Similarity	100.0%;	Pred. No. 2.6e-59;	Matches 389;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2737	AGAGCTGAGGAGTGTCAAGCTTTCATGAGTTTCAAGAGAAAGATGAGAGATTGTG	2796					
Db	390	AGAGCTGAGGAGAGTGTCAAGCTTTCATGAGTTTCAAGAGAAAGATGAGAGATTGTG	331					
QY	2797	GAAAGAGGGGAGATGCTGATTAAGATCAAGAGAAAGATGAAAGATGAAGAAGG	2856					
Db	330	GAAAGAGGGGAGATGCTGATTAAGATCAAGAGAAAGATGAAAGATGAAGAAGG	271					
QY	2857	CATGACGAGGAGATATTGTGATCTGAGAGAAAGATTTGATGAGCTTTGGAACAGCTCATG	2916					
Db	270	CATGACGAGGAGATATTGTGATCTGAGAGAAAGATTTGATGAGCTTTGGAACAGCTCATG	211					
QY	2917	TACAAGCATGGCCTTCCAAATGAAGATGATTGAGACAAAAGTCTGGTACACAAGACAGA	2976					
Db	210	TACAAGCATGGCCTTCCAAATGAAGATGATTGAGACAAAAGTCTGGTACACAAGACAGA	151					
QY	2977	CTAAGTTCTTTGTTTGTCTTTGGTATGTGCGAAAGTAGAGAGATCTGAGAGATCTCAAT	3036					
Db	150	CTAAGTTCTTTGTTTGTCTTTGGTATGTGCGAAAGTAGAGAGATCTGAGAGATCTCAAT	91					
QY	3037	TAAATFACAGGACAAATCTAAGAGATTATAGATTATATCTCCAAATTTTATAGACG	3096					
Db	90	TAAATFACAGGACAAATCTAAGAGATTATAGATTATATCTCCAAATTTTATAGACG	31					
QY	3097	GATCTAAGGAAGCATTAAGTCTTGTGAC	3125					
Db	30	GATCTAAGGAAGCATTAAGTCTTGTGAC	2					

LOCUS	CD835509	644 bp	mRNA	linear	EST 10-JUL-2003
DEFINITION	BN45.045523F011230 BN45 Brassica napus cDNA clone BN45045123, mRNA sequence.				
ACCESSION	CD835509				
VERSION	CD835509.1	GI:32517449			
KEYWORDS	EST.				
SOURCE	Brassica napus (rape)				
ORGANISM	Brassica napus Eumariotia, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, rosids, eurosids II, Brassicales, Brassicaceae, Brassica.				
REFERENCE	1 (bases 1 to 644) Genoplane.				
AUTHORS	Genoplane, a major partnership french program in plant genomics				
TITLE	Unpublished (2003)				
JOURNAL	Contact: Genoplane				
COMMENT	Genoplane 93, rue Henri Rochefort 91025 EVRY CEDEX France Tel: 33 1 69 47 54 00 Fax: 33 1 69 47 54 10 This sequence has been generated in the framework of the french plant genomics programme 'Genoplane' (http://www.genoplane.com and http://genoplane-info.infobiogen.fr). Location/Qualifiers 1..644 /organism="Brassica napus" /mol_type="mRNA" /culivar="jet neufr" /db_xref="taxon:3708" /clone="BN45045123" /tissue type="seed" /clone_lib="BN45"				
FEATURES	source				
ORIGIN	Query Match 11.8%; Score 385.8; DB 14; Length 644; Best Local Similarity 74.8%; Pred. No. 8,3e-59; Matches 533; Conservative 0; Mismatches 107; Indels 73; Gaps 1;				

[illegible]

RESULT 12	
AV525508	
LOCUS	380 bp mRNA linear EST 01-SEP-2000
DEFINITION	AV525508 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone APD25d02R 5', mRNA sequence.
ACCESSION	AV525508
VERSION	AV525508.1 GI:8665036
KEYWORDS	EST.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 380).
AUTHORS	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL	DNA Res. 7 (3), 175-180 (2000)
MEDLINE	20363093
PUBMED	10907847
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yama 1532-3, Kisarazu, Chiba 292-0812, Japan

```

FEATURES      Email: asemizukazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
SOURCE        Location/Qualifiers
               1..380
               /organism="Arabidopsis thaliana"

```

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="APD25d02R"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_id="Arabidopsis thaliana aboveground organs two to six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

```

ORIGIN

Query Match	11.6%	Score 378.4;	DB 9;	Length 380;
Best Local Similarity	99.7%	Pred. No. 2.1e-57;		
Matches 379;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0

QY	971	ATCCGGGAGAGTAAACGGCATGGTCGGGGCATTTCAAGCTTAAACAATCTCGTGGGGAGC	1030
Db	1	ATCCGGAGAGGTAAACGGCAATGGTCGGGGCATTTCAAGCTTAAACAATCTCGTGGGGAGC	60
QY	1031	AGCGTTGAGCAGAAAGATGATATACAACTTTGTGGACCCCCACCTGTATCTGGCCCTCC	1090
Db	61	AGCGTTGAGCAGAAAGATGATATACAACTTTGTGGACCCCCACCTGTATCTGGCCCTCC	120
QY	1091	TTTGGAGGAGAGATGGAATTGGCAGGCGAAGAGAGGTTCTGCTCAGACAACAGCTGTGCA	1150
Db	121	TTTGGAGGAGAGATGGAATTGGCAGGCGAAGAGAGGTTCTGCTCAGACAACAGCTGTGCA	180
QY	1151	GGAGTTTCCGACGCTGAGAGGTGATGTGGATTAATGCTTCTGAGAGAGAGATGATTTCCGA	1210
Db	181	GGAGTTTCCGACGCTGAGAGGTGATGTGGATTAATGCTTCTGAGAGAGAGATGATTTCCGA	240
QY	1211	TGCTTTGGATGATTTCTGATGACGACCTTGCACAGTGAATGATATGACTCGATGTGAATCA	1270
Db	241	TGCTTTGGATGATTTCTGATGACGACCTTGCACAGTGAATGATATGACTCGATGTGAATCA	300
QY	1271	AAAGAGCATGATCAACGAAAGCAGATTAAGTGTCAAAAAGTCTTTTGGCAGCTTGA	1330
Db	301	AAAGAGCATGATCAACGAAAGCAGATTAAGTGTCAAAAAGTCTTTTGGCAGCTTGA	360
QY	1331	TAGCTTGTCATCGAGCAGA	1350
Db	361	TAGCTTGTCATCGAGCAGA	380

RESULT 13												
LOCUS	BX547725	395 bp	DNA	linear	GSS 02-JUN-2003							
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence GK-547E01-020631, genomic survey sequence.											
ACCESSION	BX547725											
VERSION	BX547725.1	GI:32440534										
KEYWORDS	GSS.											
SOURCE	Arabidopsis thaliana (thale cress)											
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; euclidycedons; Core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.											
REFERENCE	1 Strizhov,N., Li,Y., Rosso,M., Viehoveer,P., Dekker,K., Saedler,H. and Weisshaar,B. A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines Unpublished											
TITLE	Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B. A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based genetics											
JOURNAL	2											
REFERENCE												
AUTHORS												
TITLE												

JOURNAL Unpublished
REFERENCE 3 (bases 1 to 395)
AUTHORS Strizhov N., Li Y., Rosso, M. and Weishaar, B.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2003) Weishaar B., Max-Planck-Institut fuer
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence is recovered from the left border of the T-DNA. It
indicates an insertion within the locus defined by the T-DNA. The
sequences are generated at the MPI for Plant Breeding Research in
the context of the GABI-Kat project. GABI-Kat is part of the German
Plant Genomics program designated 'GABI'. Information on line
availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES
source
1..395
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-547E01-020631"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector pAC161. The lines contain one or more T-DNA
insertions. The DNA fragment(s) resulting from the PCR
were directly sequenced to determine the genomic sequence
flanking the insertion. Sequences displaying significant
similarity to the A. thaliana nuclear genome sequence were
processed for submission. T-DNA derived sequences were
removed"

ORIGIN
Query Match 11.1%; Score 364.6; DB 29; Length 395;
Best Local Similarity 95.2%; Pred. No. 5.9e-55;
Matches 376; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 114 ACTGAATGAAATTTGGAGTCCAGATCGGAAACAGAGCGCTTTAGAGCTTAATAG 173
Db 1 ACTGAATGAAATTTGGAGTCCAGATCGGAAACAGAGCGCTTTAAAGCTTAATAG 60
QY 174 CTTCCTCATTTGCTCTCTTCTGTCAGTTATTTTCTCTCCGAGTCTGACTCACTA 233
Db 61 CTTCCTCATTTGCTCTCTTCTGTCAGTTATTTTCTCTCCGAGTCTGACTCACTA 120
QY 234 CTCTCATCTCCGGCGCTTTAAACTTAAGTCTCCGCTTACTCTGTAAGTTTCTGC 293
Db 121 CTCTCATCTCCGGCGCTTTAAACTTAAGTCTCCGCTTACTCTGTAAGTTTCTGC 180
QY 294 CTTAGAGCTCCGATCGCTCAACCGCATGCAATCTGAGTCTGATTTCTTTTCTTCCG 353
Db 181 CTTAGAGCTCCGATCGCTCAACCGCATGCAATCTGAGTCTGATTTCTTTTCTTCCG 240
QY 354 TGGAAAAATGCGCTTAATGTTCTGATTTTCAAGGTTTGTGCTATGAGTTACTTTT 413
Db 241 TGGAAAAATGCGCTTAATGTTCTGATTTTCAAGGTTTGTGCTATGAGTTACTTTT 300
QY 414 CCCATATTTTAAAGTCTTAAGTAAGTAAGTACCTGCGTCTACTGTTTGTATTTTG 473
Db 301 CCCATATTTTAAAGTCTTAAGTAAGTAAGTACCTGCAATCTAAATGTTTGTCCATTTTG 360
QY 474 TTGAGCTTTCACCGTTAGTGTGATCGAGTAT 508
Db 361 TTGAGCTTTCACCGTTAGTGTGATCGAGTAT 395

RESULT 14
AVS66465/c 619 bp mRNA linear EST 07-SEP-2000
LOCUS AVS66465 Arabidopsis thaliana green siliques Columbia Arabidopsis
DEFINITION thaliana cDNA clone SQ244b06f 3', mRNA sequence.
ACCESSION AVS66465
VERSION AVS66465.1 GI:8737917
KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
AUTHORS 1 (bases 1 to 619)
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The first laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES
source
1..619
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="SQ244b06f"
/tissue type="green siliques"
/clone_lib="Arabidopsis thaliana green siliques Columbia"
/note="Vector: pBluescriptII SK-, Site_1: EcoRI, Site_2:
XhoI"

ORIGIN
Query Match 9.2%; Score 301; DB 9; Length 619;
Best Local Similarity 72.8%; Pred. No. 1.1e-43;
Matches 513; Conservative 0; Mismatches 0; Indels 192; Gaps 2;

QY 2170 CAAGAGATGTTTAAAGAGAGTGAAGTCTGAGAGCAATCGAGCAATCGAGTGAATAC 2229
Db 619 CAAGAGATGTTTAAAGAGAGTGAAGTCTGAGAGCAATCGAGCAATCGAGTGAATAC 560
QY 2230 TTTAAGAACAGCTCTCAAAACAGAACAGACAGCCAGAGTCTTGAAGATCTTCGAA 2289
Db 559 TTTAAGAACAGCTCTCAAAACAGAACAGACAGCCAGAGTCTTGAAGATCTTCGAA 500
QY 2290 ATTATGAGGAGAGAGTGGCTGTAAGTCAAGAGATTAATGATTCGAGACAGAGACT 2349
Db 499 ATTATGAGGAGAGAGTGGCTGTAAGTCAAGAGATTAATGATTCGAGACAGAGACT 440
QY 2350 AAGATGAGCATGAACAGAACAGGAGAGATGATTTTCTAGAAAAATCAAAACTT 2409
Db 439 AAGATGAGCATGAACAGAACAGGAGAGATGATTTTCTAGAAAAATCAAAACTT 412
QY 2410 GACATTTTGTATTAACCTAGTCAATTTTGATTAATGTCACAAAAAACTGT 2469
Db 411 ----- 412
QY 2470 GGTGTTGAAGATGATGACAGACAGGTTTTCATGATTCATCAACAGATTCAT 2529
Db 411 -----AGATGATGACAGACAGGTTTTCATGATTCATCAACAGATTCAT 362
QY 2530 GAAAGAGAGACCCAAAGAGAGAAATTTGAGATGTTGAGAGACAGAGAGTCCCAAG 2589
Db 361 GAAAGAGAGACCCAAAGAGAGAAATTTGAGATGTTGAGAGACAGAGAGTCCCAAG 302
QY 2590 GTTGTGGCCAGACAGACAGAACTTATTCCTCTTACATATGATGATTCGCAAGAGG 2649
Db 301 GTTGTGGCCAGACAGACAGAACTTATTCCTCTTACATATGATGATTCGCAAGAGG 245
QY 2650 TATATGATCACTAATATATCCCTGCGTTTGTGTTTCAACCTAAGATTAAT 2709
Db 244 ----- 245
QY 2710 GAATTAATCCGTTTGAATTTCTTCCAGAGCTGAGAAAGTCAAGCTTCATCGATT 2769

Db 244 -----AGACCTGAGGAGTGTCAAGCTTCATCGAGTTT 212

QY 2770 CAAGAGAAAGATGATGAGAGTTTGTGGAGAGAGGGAGATGCTGATTAAGATCAAGAG 2829

Db 211 CAAGAGAAAGATGAGAGAGTTTGTGGAGAGAGGGAGATGCTGATTAAGATCAAGAG 152

QY 2830 AAGAAGATGGAAGATGAGAGAGGAGATGATGAGAGAGATATT 2874

Db 151 AAGAAGATGGAAGATGAGAGAGGAGATGATGAGAGATATT 107

RESULT 15

AQ010650 290 bp DNA linear GSS 29-MAY-1998

LOCUS F27C8TRC IGF Arabidopsis thaliana genomic clone F27C8, genomic

DEFINITION survey sequence.

ACCESSION AQ010650

VERSION AQ010650.1 GI:3165927

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE

AUTHORS Rounsley, S.D., Suh, B.J., Wible, C., Golden, K., Shatsman, S., Choi, P., Yu, K., Akinteyo, B., Shen, K., Geonasekaram, S., Miltner, J., Adams, M.D. and Venter, J.C.

A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 4

Unpublished (1998)

Other GSSs: F27C8TRC

Contact: Steve Rounsley

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: rounsley@tigr.org

Seq primer: M13 Reverse

Class: BAC ends

High quality sequence stop: 290.

TITLE

JOURNAL

COMMENT

Location/Qualifiers

1..290

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="F27C8"

/sex="hermaphrodite"

/clone_1lb="IGF"

/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;

Produced by Thomas Altmann"

FEATURES

source

1..290

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="F27C8"

/sex="hermaphrodite"

/clone_1lb="IGF"

/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;

Produced by Thomas Altmann"

ORIGIN

Query Match 8.9%; Score 290; DB 28; Length 290;

Best Local Similarity 100.0%; Pred. No. 1.4e-41;

Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1668 TTCTGCTTTTACTTTCTTTAATTTTCTCTGCACTTCTAGATCTTGAAGTTTACATT 1727

Db 1 TTCTGCTTTTACTTTCTTTAATTTTCTCTGCACTTCTAGATCTTGAAGTTTACATT 60

QY 1728 GTAGTGCTGGGCAATGGGCAACGAAGCTGCTGAATTAATTGCAAGTATGAGGCTCT 1787

Db 61 GTAGTGCTGGGCAATGGGCAACGAAGCTGCTGAATTAATTGCAAGTATGAGGCTCT 120

QY 1788 TAGAGCAGGCATTCCTATGCTCCACAGGGCCATGCTGGATGAGTGTCTGATGTTGA 1847

Db 121 TAGAGCAGGCATTCCTATGCTCCACAGGGCCATGCTGGATGAGTGTCTGATGTTGA 180

QY 1848 GAGCAGTCCACTGCTATTGAGGCGGAAGCCTCCACCGGAGTTAGCTGAGATGG 1907

Db 181 GAGCAGTCCACTGCTATTGAGGCGGAAGCCTCCACCGGAGTTAGCTGAGATGG 240

QY 1908 GTTAGATGAAATTCCTGGGGGTGAGAAGCGCAGTATGTTTCTGGAGGTG 1957

Db 241 GTTAGATGAAATTCCTGGGGGTGAGAAGCGCAGTATGTTTCTGGAGGTG 290

Search completed: April 8, 2004, 16:24:24

Job time : 5740.02 secs

XX	Beclin C,	Elmayer T,	Vaucheret H;
PI	WPI,	2001-159529/16.	
DR	New SGS3 gene from	Arabidopsis thaliana,	useful for increasing virus resistance in plants and, when inhibited, for increasing transgene expression.
XX	Claim 1,	Page 31-32,	36pp; French.
PS	The present sequence	represents the genomic sequence of the Arabidopsis thaliana SGS3 gene. The SGS3 gene is essential for post-transcriptional inactivation (degradation of RNA) and for resistance to viruses. Overexpression of SGS3 results in plants with increased resistance to viruses, while inactivation of SGS3 in transgenic plants (e.g. by expressing antisense RNA, by mutation or by homologous recombination) increases the level of the transgene product. This product may e.g. impart resistance (to herbicide, insects or pathogens), alter contents of essential fatty acids or proteins, or is pharmaceutically active, e.g. an immunoglobulin or interferon	
CC	Sequence 3275 BP,	956 A;	561 C; 804 G; 954 T; 0 U; 0 Other;
SQ	Query Match	100.0%;	Score 3275; DB 4; Length 3275;
	Best Local Similarity	100.0%;	Pred. No. 0;
	Matches 3275;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	GACAAACAACAATAATTAGCAATTCATGTTCTGTAGCATTAATAATTAGTGGGAACA	60
DB	1	GACAAACAACAATAATTAGCAATTCATGTTCTGTAGCATTAATAATTAGTGGGAACA	60
QY	61	TTAAGTTAAGCGAAAAAGAAAAAGTAGTCAAAAATGAAAAACAATCAAATCTGAAT	120
DB	61	TTAAGTTAAGCGAAAAAGAAAAAGTAGTCAAAAATGAAAAACAATCAAATCTGAAT	120
QY	121	GAAAAATTGGAGTCCAGAATCGGAAAAACGAGGCCGTTTGAAGCTTAATAAGCTTCTC	180
DB	121	GAAAAATTGGAGTCCAGAATCGGAAAAACGAGGCCGTTTGAAGCTTAATAAGCTTCTC	180
QY	181	AATTTGTCCTCTCTTGTCAGTTTATTTCTTCTTCGCGAGTCTGACACACTACTCAG	240
DB	181	AATTTGTCCTCTCTTGTCAGTTTATTTCTTCTTCGCGAGTCTGACACACTACTCAG	240
QY	241	TCCTCCGCGCTTTAAACTTAAGCTTCGCGTGTACTCTGAAGTTTTGCCTTAGAG	300
DB	241	TCCTCCGCGCTTTAAACTTAAGCTTCGCGTGTACTCTGAAGTTTTGCCTTAGAG	300
QY	301	CCTCGAGTCGCGCTCACGCGCATGCAATCTGTGCTGCAATTTCTTTCTGCGCGAAAA	360
DB	301	CCTCGAGTCGCGCTCACGCGCATGCAATCTGTGCTGCAATTTCTTTCTGCGCGAAAA	360
QY	361	ATTGCGCTTAATGTTCTCGATTTTCGAAGGTTTTGTCTATGAGGTACTTTTTCCCTATA	420
DB	361	ATTGCGCTTAATGTTCTCGATTTTCGAAGGTTTTGTCTATGAGGTACTTTTTCCCTATA	420
QY	421	TTTATAAGTCTTAAAGTAAACATACCTCCGCTCTACTGTTTTGTCAATTTTGTGTGCT	480
DB	421	TTTATAAGTCTTAAAGTAAACATACCTCCGCTCTACTGTTTTGTCAATTTTGTGTGCT	480
QY	481	TTCAACCGTTTAGTGCATCGAGATATTTACTGTGAAAAATCCTTGTTTTTGGTTTT	540
DB	481	TTCAACCGTTTAGTGCATCGAGATATTTACTGTGAAAAATCCTTGTTTTTGGTTTT	540
QY	541	TTGTTTCAATAAATCGAATTGATCTACCTTTTGCTTTGATGTTTTGTTTGAAGCT	600
DB	541	TTGTTTCAATAAATCGAATTGATCTACCTTTTGCTTTGATGTTTTGTTTGAAGCT	600
QY	601	ATGCGGTGTGGCTGTATTAACCTCAAGTCAAGTGTGAAATTTGAAATTTGGTAGTG	660
DB	601	ATGCGGTGTGGCTGTATTAACCTCAAGTCAAGTGTGAAATTTGAAATTTGGTAGTG	660
QY	661	ACTGAGGTTTCTTTGCTGCTATAGTTGTAAAAATGATGTTCTAAGGCTGTGCCAAGT	720

Db	661	ACTGTTGGGTTCTTTGGCTGATAGTTGTGTAATAATGATGTTCTAGGGCTGCTCCAAATG	720
QY	721	CTAAGAAAGAACGTTTCAGGGTGGTTATAGAGCCCTGAGGTTGAACATGTGGTTCAAGGTT	780
Db	721	CTAAGAAAGAACGTTTCAGGGTGGTTATAGAGCCCTGAGGTTGAACATGTGGTTCAAGGTT	780
QY	781	TGGCAGGGACGAGACTGGCTTCTTCACAAGATGATGAGAGAGAGTGGAGGTCATTTCGA	840
Db	781	TGGCAGGGACGAGACTGGCTTCTTCACAAGATGATGAGAGAGAGTGGAGGTCATTTCGA	840
QY	841	AGAAGAACAAAGAACAAACCCAGAGAAACATTCCTGGAAAAAATTGGGTTTCTCAGAAATTCGA	900
Db	841	AGAAGAACAAAGAACAAACCCAGAGAAACATTCCTGGAAAAAATTGGGTTTCTCAGAAATTCGA	900
QY	901	ATCCTCTTAGAGCTTGGGGTGGTCAGCAGCAAGGAGAGGTAGCAACGTATCTGGAGAG	960
Db	901	ATCCTCTTAGAGCTTGGGGTGGTCAGCAGCAAGGAGAGGTAGCAACGTATCTGGAGAG	960
QY	961	GAACAAATGTATCCGGGAGAGGTAAACGGCAATGTGTCGGGGCATTCAGGCTAAATATCTG	1020
Db	961	GAACAAATGTATCCGGGAGAGGTAAACGGCAATGTGTCGGGGCATTCAGGCTAAATATCTG	1020
QY	1021	GTGCGGGACGAGCGGTTGAGCAGAAAGTATGATTAACAATTGTGTGGCAACCCCACTGTAT	1080
Db	1021	GTGCGGGACGAGCGGTTGAGCAGAAAGTATGATTAACAATTGTGTGGCAACCCCACTGTAT	1080
QY	1081	CTGCGCCCTCTTTGGAAAGAGAGATGGAATTGGCAGGCAAGAGAGGTTCTGCTCAGACA	1140
Db	1081	CTGCGCCCTCTTTGGAAAGAGAGATGGAATTGGCAGGCAAGAGAGGTTCTGCTCAGACA	1140
QY	1141	CAGCTGTGACAGAGATTTCTCAACCTGAGAGATGATGTGATTAATGCTTCTAGAGAAAGAG	1200
Db	1141	CAGCTGTGACAGAGATTTCTCAACCTGAGAGATGATGTGATTAATGCTTCTAGAGAAAGAG	1200
QY	1201	ATGATTTCCGATGCTTTGGATGATATTCGATGACGACCTTTCAGAGATGATTAATGACTCGG	1260
Db	1201	ATGATTTCCGATGCTTTGGATGATATTCGATGACGACCTTTCAGAGATGATTAATGACTCGG	1260
QY	1261	ATGAGAGCAAAAGAGCCATGATCAAGAAAGCAGATTAAGTGGTTCAAAAAGTCTCTTG	1320
Db	1261	ATGAGAGCAAAAGAGCCATGATCAAGAAAGCAGATTAAGTGGTTCAAAAAGTCTCTTG	1320
QY	1321	GCAGCTTGATAGCTTGTGATTCGATTCGACGAGATTAATGAACACACAGAGCAGTGGCATTTGC	1380
Db	1321	GCAGCTTGATAGCTTGTGATTCGATTCGACGAGATTAATGAACACACAGAGCAGTGGCATTTGC	1380
QY	1381	CAGCTGTGACAAAGAGACCTGGTGTCCATGATGTGATTAACCTGACACCCTCTTACTACTC	1440
Db	1381	CAGCTGTGACAAAGAGACCTGGTGTCCATGATGTGATTAACCTGACACCCTCTTACTACTC	1440
QY	1441	ATGCGAGAGCAAAAGAGAGCTAGGCGAGGATTAAGCTCCATAGAGATTTGGCTGAAGTTTAG	1500
Db	1441	ATGCGAGAGCAAAAGAGAGCTAGGCGAGGATTAAGCTCCATAGAGATTTGGCTGAAGTTTAG	1500
QY	1501	AAAAGAGATCTACAGATGAGAGGCCATCTGTCAATCTTGTGGTGAATTTATGGCAGT	1560
Db	1501	AAAAGAGATCTACAGATGAGAGGCCATCTGTCAATCTTGTGGTGAATTTATGGCAGT	1560
QY	1561	GGAAGGGTTGGGTGAGAGATGAAGAAAGATTAATGAATTTGCTGGCCCTCCAAATGTCATCA	1620
Db	1561	GGAAGGGTTGGGTGAGAGATGAAGAAAGATTAATGAATTTGCTGGCCCTCCAAATGTCATCA	1620
QY	1621	TCATGAATTAAGCTGAGATGAGAGCAGATTAACGATTAAGTGAATTTCTTGTCTTTTAC	1680
Db	1621	TCATGAATTAAGCTGAGATGAGAGCAGATTAACGATTAAGTGAATTTCTTGTCTTTTAC	1680
QY	1681	TTCTTTAATTTTCTCTTGATTTCTAATCTTAAGATGTTACATTTGTAGTGGCTGGC	1740
Db	1681	TTCTTTAATTTTCTCTTGATTTCTAATCTTAAGATGTTACATTTGTAGTGGCTGGC	1740
QY	1741	ATGGGCAACAGAGCTGTGTAATTAATTCGACAGATTAAGGCTCTTAAGCAGCGCAT	1800

Dp	1741	ATGGGCAACAAAGAGCTGCTGGAAATTA	CTCGACAAGATATGAGGCTCTTAGAGCAGCCAT	1801
Qy	1801	TCCTATGCTCCACAGGGCCATTCGTGGAGT	TGAGTGTCTGATGTTTGAGACAGTGCAC	1860
Dp	1801	TCCTATGCTCCACAGGGCCATTCGTGGAGT	GTGTCTGATGTTTGAGACAGTGCAC	1860
Qy	1861	GGCATATTTGAGAGCCGAACGCGCTCCAC	CCGGGAGTTAGCTGAGATGAGGGCTTGATGAA	1920
Dp	1861	GGCATATTTGAGAGCCGAACGCGCTCCAC	CCGGGAGTTAGCTGAGATGAGGGCTTGATGAA	1920
Qy	1921	GCTTGGGGCTCAGAAAGCGCAGTATGTTT	CTGGAAGGTGTTCCCACTGATAGCTTCCT	1980
Dp	1921	GCTTGGGGCTCAGAAAGCGCAGTATGTTT	CTGGAAGGTGTTCCCACTGATAGCTTCCT	1980
Qy	1991	GCAACGAAGCAGATCTGACATATTCATCA	CACTCTCAAGGTTCTCTCCCAAGAA	2040
Dp	1991	GCAACGAAGCAGATCTGACATATTCATCA	CACTCTCAAGGTTCTCTCCCAAGAA	2040
Qy	2041	AATTTGATATATGCTTTTAAGTTTGTCA	TGGAATTTAAAGTTTGTGTGCTGGTTAA	2100
Dp	2041	AATTTGATATATGCTTTTAAGTTTGTCA	TGGAATTTAAAGTTTGTGTGCTGGTTAA	2100
Qy	2101	TGCATCTGTATGTAATATCTATGATTC	ATTAGGCAAAACAAGGCTGAATTCGAGTTG	2160
Dp	2101	TGCATCTGTATGTAATATCTATGATTC	ATTAGGCAAAACAAGGCTGAATTCGAGTTG	2160
Qy	2161	AAATCATACCAAGAGATGTTGTAAAGAG	CTGAGCAGATCTCTGAGACAAATCAGCAG	2220
Dp	2161	AAATCATACCAAGAGATGTTGTAAAGAG	CTGAGCAGATCTCTGAGACAAATCAGCAG	2220
Qy	2221	CTGAACATCTTTAAGAACAAAGCTCTCA	AAACAAGAACGCAAGGCTGCTTGAAGAA	2280
Dp	2221	CTGAACATCTTTAAGAACAAAGCTCTCA	AAACAAGAACGCAAGGCTGCTTGAAGAA	2280
Qy	2281	TCTCTGAAATTTATATGAGCGAAGAG	CTGCGTGAACCTCGACAGAGATTAATCG	2340
Dp	2281	TCTCTGAAATTTATATGAGCGAAGAG	CTGCGTGAACCTCGACAGAGATTAATCG	2340
Qy	2341	CAGAGAACCTAAGATGACAGCATGAAC	CAGAACAGGAAAGATATGATTTTCTTGAA	2400
Dp	2341	CAGAGAACCTAAGATGACAGCATGAAC	CAGAACAGGAAAGATATGATTTTCTTGAA	2400
Qy	2401	CACAAACTTGAACATTTGTATTA	CTACTGATTCACATTTTGTATATATGTC	2460
Dp	2401	CACAAACTTGAACATTTGTATTA	CTACTGATTCACATTTTGTATATATGTC	2460
Qy	2461	AAAACCTGTGGTGTGAAGATGATGATG	ATGATGATGATGATGATGATGATGATG	2520
Dp	2461	AAAACCTGTGGTGTGAAGATGATGATG	ATGATGATGATGATGATGATGATGATG	2520
Qy	2521	CAGATCCATGAAGAAAGACGCAAAAG	GAGAGATTTGAGATGTTGACAGACAGAA	2580
Dp	2521	CAGATCCATGAAGAAAGACGCAAAAG	GAGAGATTTGAGATGTTGACAGACAGAA	2580
Qy	2581	CGTGCCAAAGGTTGTGGCCACAGACAG	ACAGAACTTAATCCCTCAGCAATGACG	2640
Dp	2581	CGTGCCAAAGGTTGTGGCCACAGACAG	ACAGAACTTAATCCCTCAGCAATGACG	2640
Qy	2641	CGAAAGAGGTATATGTAATTA	CAATTAATCCCTCGGCGTTTGTGTTT	2700
Dp	2641	CGAAAGAGGTATATGTAATTA	CAATTAATCCCTCGGCGTTTGTGTTT	2700
Qy	2701	AGAGTAATCTGAATTAATTCGGT	TTTTGATCTTTTCGACAGAGCTGAGAA	2760
Dp	2701	AGAGTAATCTGAATTAATTCGGT	TTTTGATCTTTTCGACAGAGAGTGTCA	2760
Qy	2761	ATCGAGTTTCAAGGAAAGAGATGAGAG	AGTTTGTGAAAGAGAGGAGATCTGAT	2820
Dp	2761	ATCGAGTTTCAAGGAAAGAGATGAGAG	AGTTTGTGAAAGAGAGGAGATCTGAT	2820
Qy	2821	GATCAAGAGAAAGATGGAAGACATGA	AGAAAGGCAATCAAGAGATTTTGATCTG	2880
Dp	2821	GATCAAGAGAAAGATGGAAGACATGA	AGAAAGGCAATCAAGAGATTTTGATCTG	2880

QY	2881	GAGAAAGAAATTGTAAGAGGCTTTGGAAACAGCTCATGTATCAAGCATGCGCTTCACAATGAA	2940
Db	2881	GAGAAAGAAATTGTAAGAGGCTTTGGAAACAGCTCATGTATCAAGCATGCGCTTCACAATGAA	2940
QY	2941	GATGATTGAGACAAAGCTGTGTAACACAGCAAGACTAAAGTTTCCTTGTGTTGCTTTTG	3000
Db	2941	GATGATTGAGACAAAGCTGTGTAACACAGCAAGACTAAAGTTTCCTTGTGTTGCTTTTG	3000
QY	3001	GTAATGTGGAAAGTATGAGAGATCTGAGAGACTCCATTTTAAATCTATGAGACAAATTTAAGA	3060
Db	3001	GTAATGTGGAAAGTATGAGAGATCTGAGAGACTCCATTTTAAATCTATGAGACAAATTTAAGA	3060
QY	3061	GATTATGATATTATTCCTCCAAATTTTTTAGTAGACGGATCTAAGAAAGCATTAAGTCTT	3120
Db	3061	GATTATGATATTATTCCTCCAAATTTTTTAGTAGACGGATCTAAGAAAGCATTAAGTCTT	3120
QY	3121	GTGACTAAACCAAGTTTCCTTAGTATTGTTTTTTTTTGGTAAAAATTCATATGAAG	3180
Db	3121	GTGACTAAACCAAGTTTCCTTAGTATTGTTTTTTTTTGGTAAAAATTCATATGAAG	3180
QY	3181	TTAGACATATTACCAACGTCAGAGTGATATACGAATGCGAATCAAAATCATGTTTTT	3240
Db	3181	TTAGACATATTACCAACGTCAGAGTGATATACGAATGCGAATCAAAATCATGTTTTT	3240
QY	3241	AGAAATTTATATCTACAAATATATATGGGTACAAT	3275
Db	3241	AGAAATTTATATCTACAAATATATATGGGTACAAT	3275
RESULT 2			
ID	AAF25374	standard; cDNA; 1878 BP.	
AC	AAF25374;		
XX			
DT	15-MAY-2001	(first entry)	
XX			
DE		Nucleotide sequence of the Arabidopsis SGS3 polypeptide.	
XX			
KM	SGS3 gene; post-transcriptional inactivation; RNA degradation;		
KW	viral resistance; resistance; fatty acid content; protein content; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	MO200105951-A2.		
XX			
PD	25-JAN-2001.		
XX			
PF	13-JUL-2000; 2000WO-FR002052.		
XX			
PR	16-JUL-1999; 99FR-00009417.		
XX	26-JAN-2000; 2000FR-00001006.		
XX			
PA	(AVET) AVENTIS CROSCIENCE SA.		
XX	PA (INRG) INST NAT RECH AGRONOMIQUE.		
XX			
PI	Beclin C, Elmayan T, Vaucheret H;		
XX			
DR	WP1; 2001-159529/16.		
XX	P-PSDB; AAB31798.		
XX			
PT	New SGS3 gene from Arabidopsis thaliana, useful for increasing virus		
XX	resistance in plants and, when inhibited, for increasing transgene		
XX	expression.		
XX			
XX	Claim 1; 32-35; 36pp; French.		

CC The present sequence encodes the Arabidopsis thaliana SGS3 polypeptide.
 CC The SGS3 gene is essential for post-transcriptional inactivation
 CC (degradation of RNA) and for resistance to viruses. Overexpression of
 CC SGS3 results in plants with increased resistance to viruses, while
 CC inactivation of SGS3 in transgenic plants (e.g. by expressing antisense
 CC RNA, by mutation or by homologous recombination) increases the level of
 CC the transgene product. This product may e.g. impart resistance (to
 CC herbicide, insects or pathogens), alter contents of essential fatty acids
 CC or proteins, or is pharmaceutically active, e.g. an immunoglobulin or
 CC interferon

XX Sequence 1878 BP; 575 A; 332 C; 554 G; 417 T; 0 U; 0 Other;

Query Match 44.6%; Score 1462; DB 4; Length 1878;

Best Local Similarity 83.3%; Pred. No. 0;

Matches 1878; Conservative 0; Mismatches 0; Indels 376; Gaps 4;

QY 696 ATGAGTTCTAGGGCTGGTCCAAATGCTTAAGAAAGAACGTTGAGGTTGTTAAGGCTT 755
 DB 1 ATGAGTTCTAGGGCTGGTCCAAATGCTTAAGAAAGAACGTTGAGGTTGTTAAGGCTT 60
 QY 756 GAGTTGAACAGTTGGTTCAAGTTTGGCAGGAGCAGACTGGCTTCTTCAAGATGAT 815
 DB 61 GAGTTGAACAGTTGGTTCAAGTTTGGCAGGAGCAGACTGGCTTCTTCAAGATGAT 120
 QY 816 GGAGGAGAGTGGAGGTCATTTCAGAAAGAACAGAACAGAACAGAACAGAACAGTCTTGG 875
 DB 121 GGAGGAGAGTGGAGGTCATTTCAGAAAGAACAGAACAGAACAGAACAGAACAGTCTTGG 180
 QY 876 AAAAATTGGGTTTCTCAGAAATTCGATCTCTAGAGCTTGGGTTGCTCAGCAGAAAGG 935
 DB 181 AAAAATTGGGTTTCTCAGAAATTCGATCTCTAGAGCTTGGGTTGCTCAGCAGAAAGG 240
 QY 936 AAGAGTGAACAGTTTCTGGGAGAGAAACATGATCCGGAGAGGTAAACCGCATTGGT 995
 DB 241 AAGAGTGAACAGTTTCTGGGAGAGAAACATGATCCGGAGAGGTAAACCGCATTGGT 300
 QY 996 CGGGGATTCAGAGCTAATATCTGGTCGGGAGCAGAGCTTGGAGAAAGTATGATTAAC 1055
 DB 301 CGGGGATTCAGAGCTAATATCTGGTCGGGAGCAGAGCTTGGAGAAAGTATGATTAAC 360
 QY 1056 AACTTTGTGGACCCCACTGTATTCGCTCTTTGGAGAGAGATGAAATTGGCAG 1115
 DB 361 AACTTTGTGGACCCCACTGTATTCGCTCTTTGGAGAGAGATGAAATTGGCAG 420
 QY 1116 GCAAGAGAGGTCTGCTCAGACACAGCTGTGCAAGAGGTTCTGACGTGAGAGTAT 1175
 DB 421 GCAAGAGAGGTCTGCTCAGACACAGCTGTGCAAGAGGTTCTGACGTGAGAGTAT 480
 QY 1176 GTGATTAATGCTTCTGAGAGAGAGATGATTCGATGCTTTGGATGATCTGATGACGAC 1235
 DB 481 GTGATTAATGCTTCTGAGAGAGAGATGATTCGATGCTTTGGATGATCTGATGACGAC 540
 QY 1236 CTTCGAGATGATTAATGATCTCGATGTGATCAAAAGAGCCATGATCAAGAAAGCAG 1295
 DB 541 CTTCGAGATGATTAATGATCTCGATGTGATCAAAAGAGCCATGATCAAGAAAGCAG 600
 QY 1296 AATAAGTGTCAAAAGTCTTTGGCAGCTTGGATGATGATGATGATGATGATGATGAT 1355
 DB 601 AATAAGTGTCAAAAGTCTTTGGCAGCTTGGATGATGATGATGATGATGATGATGAT 660
 QY 1356 GAACCAAG 1415
 DB 661 GAACCAAG 720
 QY 1416 TATTAACCTCAGCTCTTACTAGCTCATGCGAGAGCAAAAGAGAGAGAGAGAGAGAG 1475
 DB 721 TATTAACCTCAGCTCTTACTAGCTCATGCGAGAGCAAAAGAGAGAGAGAGAGAGAG 780
 QY 1476 CATAGAGAAATGGCTGAAGTTTGAAGAGATCTCAATGAGAGAGAGAGAGAGAGAGAG 1535
 DB 781 CATAGAGAAATGGCTGAAGTTTGAAGAGATCTCAATGAGAGAGAGAGAGAGAGAGAG 840

QY 1536 CCTTGGTGAGATTTATGAGGAGTGAAGAGGTTTGGTGAGATGATAAGAGATTTATGA 1595
 DB 841 CCTTGGTGAGATTTATGAGGAGTGAAGAGGTTTGGTGAGATGATAAGAGATTTATGA 900
 QY 1596 ATTGTCTGGCTCCAAATGTCATCATGAAATTAATTAAGCTGATTAAGAGCAGTAACAT 1655
 DB 901 ATTGTCTGGCTCCAAATGTCATCATGAAATTAATTAAGCTGATTAAGAGCAGTAACAT 960
 QY 1656 AAGGTGGAATTTCTGTCTTTACTTCTTAATTTTCTCTGCAATCTACTGATCTTA 1715
 DB 961 A----- 961
 QY 1716 GAATGTTACATTTAGTGGCTGGCATGGCAGCAACCAAGAGCTGTGAAATCTTGACAA 1775
 DB 962 -----AGTGGCTGGCATGGCAGCAACCAAGAGCTGTGAAATCTTGACAA 1007
 QY 1776 GTATGAGCTCTTAAGACACGCAATTTCTTAATGTCACAGGAGCATCGTGGATGATGT 1835
 DB 1008 GTATGAGCTCTTAAGACACGCAATTTCTTAATGTCACAGGAGCATCGTGGATGATGT 1067
 QY 1836 TCTGATTTTGAAGACAGTGCACCTGGCTATTGAGAGCCCAACGCTTCCACCGGGAGTT 1895
 DB 1068 TCTGATTTTGAAGACAGTGCACCTGGCTATTGAGAGCCCAACGCTTCCACCGGGAGTT 1127
 QY 1896 ACCTGAGATGGGTTAGATAGAAATGGCTGGGAGTCAAGAGCCAGATGTTTCTGAGAG 1955
 DB 1128 ACCTGAGATGGGTTAGATAGAAATGGCTGGGAGTCAAGAGCCAGATGTTTCTGAGAG 1187
 QY 1956 TGTTCGCAACTGTATGAGCTTCTTGAACAGAGCAAGATCTGACATATTCATCAACA 2015
 DB 1188 TGTTCGCAACTGTATGAGCTTCTTGAACAGAGCAAGATCTGACATATTCATCAACA 1247
 QY 2016 CTCTGAAAGTTCTCTCCCCCAAGAAATTTGATATATGCTTTAGTTTTCATTTGAT 2075
 DB 1248 CTCTCA----- 1253
 QY 2076 TTAAGTTTGTGGTCGGTTAATGATCTGTTATGATATCTATGATTCATTAGG 2135
 DB 1254 -----AGG 1256
 QY 2136 CAAAACAAAGCTGAATTCAGTTGAATCATACCAAGAGATGTTGTAAGAGACTGAG 2195
 DB 1257 CAAAACAAAGCTGAATTCAGTTGAATCATACCAAGAGATGTTGTAAGAGACTGAG 1316
 QY 2196 GCAATCTCTGAGCAATCAGCAGCTGAACCTAATTAAGAAACAAAGCTCTCAAAACAG 2255
 DB 1317 GCAATCTCTGAGCAATCAGCAGCTGAACCTAATTAAGAAACAAAGCTCTCAAAACAG 1376
 QY 2256 CAAGCAGCAGAGTCTTGAAGATCTCTGAAATTTAGCGAGAGAGCTGCTGTAAC 2315
 DB 1377 CAAGCAGCAGAGTCTTGAAGATCTCTGAAATTTAGCGAGAGAGCTGCTGTAAC 1436
 QY 2316 TGCAGAGATTAATCGATCTGAGACAGAGAACTAAGATGCAAGCATGAACAGAGAG 2375
 DB 1437 TGCAGAGATTAATCGATCTGAGACAGAGAACTAAGATGCAAGCATGAACAGAGAG 1496
 QY 2376 AGAGTATGATTTTCTTAAGAAATCAAAACTTGAACATTTGTATTACTACTGATTC 2435
 DB 1497 AG----- 1498
 QY 2436 CATTTTGTATTAATTTGTCACAAACAAAACCTGTGTGTTGAAGATGATGACACAGA 2495
 DB 1499 -----AGATGATGACACAGA 1514
 QY 2496 CAGTTTTTTCATGATTCATCAATCAACAGATCCATGAAGAGAGAGAGAGAGAGAGAG 2555
 DB 1515 CAGTTTTTTCATGATTCATCAATCAACAGATCCATGAAGAGAGAGAGAGAGAGAGAG 1574
 QY 2556 TTTGAGATTTTGAAGAT 2615
 DB 1575 TTTGAGATTTTGAAGAT 1634
 QY 2616 TAATCCCTCTAGCAATGACGATTCGCCAAGAGATATGTAATTAATCAATATCCCT 2675

```

Db      1635 TAATCCCTAGCAATGACGATTGCCGAAG----- 1665
Qy      2676 CTGGCGTTTGTGTTTCAAACTAAGAGTAAGTATTCGCGTTTGTATTCCTTCG 1735
Db      1666 ----- 1665
Qy      2736 CAGAGCTGAGGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGGATTTGT 2795
Db      1666 -AGAGCTGAGGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGGATTTGT 1724
Qy      2796 GGAAGAGAGGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGGATTTGT 2855
Db      1725 GGAAGAGAGGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGGATTTGT 1784
Qy      2856 GCATCAGGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGGATTTGT 2915
Db      1785 GCATCAGGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGGATTTGT 1844
Qy      2916 GTACAGCATGCGCTTCACAAATGAGATGATGA 2949
Db      1845 GTACAGCATGCGCTTCACAAATGAGATGATGA 1878

RESULT 3
ABX81661
ID      ABX81661 standard; cDNA; 272 BP.
XX      ABX81661;
AC      ABX81661;
XX      24-APR-2003 (first entry)
DT      24-APR-2003 (first entry)
XX      24-APR-2003 (first entry)
DE      Corn ear-derived polynucleotide (cpd) #121.
XX      Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
XX      structural gene; functional gene; regulatory gene;
XX      corn ear-specific profile; gene transcription; gene expression;
XX      hybrid plant; desirable trait expression; plant breeding program;
XX      inheritance; desirable characteristics; growth; development;
XX      disease resistance; environmental adaptability; quality; yield;
XX      multigene trait; plant; gene; ss.
XX      Zea mays.
OS      Zea mays.
XX      US6476212-B1.
XX      05-NOV-2002.
XX      14-MAY-1999; 99US-00313294.
XX      26-MAY-1998; 98US-0086722P.
XX      (INCY-) INCYTE GENOMICS INC.
XX      (INCY-) INCYTE GENOMICS INC.
XX      LaJuvdi RV, Ito LY, Sherman BK;
XX      WPI; 2003-208840/20.
XX      Novel purified corn-ear derived polynucleotide useful as hybridization
XX      probe for detecting polynucleotide in sample, and for identifying,
XX      evaluating, and altering desired characteristics associated with growth,
XX      development.
XX      Example; SEQ ID NO 121; 390pp; English.
XX      The present invention relates to the isolation of corn ear-derived
XX      polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
XX      and SATMON023. Some of the cdps uniquely identify structural, functional,
XX      and regulatory genes of corn ear. The polynucleotides sequences are
XX      useful for detecting cdps in a sample, for producing a corn ear-specific
XX      profile of gene transcription, for detecting altered gene expression in
XX      inbred or hybrid plants, and for screening several molecules for specific
XX      binding to the polynucleotide. The cdps are useful to identify, isolate,

```

```

CC      or extend identical or related corn-ear nucleic acid sequences from DNA
CC      libraries, and in nucleic acid amplification or hybridization techniques
CC      to follow the expression of desirable traits through plant breeding
CC      program. Preferably, the cdps are used to identify, evaluate, alter, or
CC      follow the inheritance of desirable characteristics associated with growth
CC      and development, disease resistance, environmental adaptability, quality,
CC      and yield of corn. The cdps are also useful as molecular markers for
CC      studying inheritance and multigene traits in a plant breeding program.
CC      The cdps are useful for producing purified corn-ear polypeptides by
CC      recombinant techniques. They are also useful in diagnostic assays to
CC      detect or confirm conditions or diseases associated with abnormal levels
CC      of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC      polynucleotides (cdps) of the invention. Note: The sequence data for this
CC      patent did not form part of the printed specification, but was obtained
CC      in electronic format directly from the USPTO web site at
CC      seqdata.uspto.gov/psipdb/entry.html
XX      SQ
XX      Sequence 272 BP; 79 A; 52 C; 68 G; 67 T; 0 U; 6 Other;
XX      Query Match 2.3%; Score 74.4; DB 7; Length 272;
XX      Best Local Similarity 58.7%; Pred. No. 7.8e-08;
XX      Matches 142; Conservative 0; Mismatches 97; Indels 3; Gaps 1;

Qy      1777 TATGAGCTCTTAGAGCAGCCATTCTTATGTCACAGGCGCATGTCGATGATGTT 1836
Db      2 TATGAGCAGAGTMAAGCAGTCATGCTATGTCCTCGGGCAGCGTGTATGAGCGTN 61
Qy      1837 CTGATGTTTGAAGCAGTCGCCACTGGCTATTTTGAAGCGCGAAAGCCTCCACCGGAGTTA 1896
Db      62 TTAATATTTGAAAGCTCAGCTGTGGCTTACATGGAAGCTGAACGTCTGATTAACACTTT 121
Qy      1897 GCTGAGATGGGTTGATGATGAATGCTGG---GGTCAGAACGCGATATGTTTCTGGA 1953
Db      122 GTTAANCAAGTACAGACAGAAATTCATGCACTTACGAGAGTTGATTTGGCTGAT 181
Qy      1954 GGTGTTGCGCAACTGATGCTTCTTGCAGAGAGCAAGATCTGGACATATTCATCAA 2013
Db      182 GGGAAAGCCNCCTATATGTTTCTTGAACAACAAAGAGATATGAGCCATTTAACAAG 241
Qy      2014 CA 2015
Db      242 CA 243

RESULT 4
ABX82003
ID      ABX82003 standard; cDNA; 274 BP.
XX      ABX82003;
AC      ABX82003;
XX      24-APR-2003 (first entry)
DT      24-APR-2003 (first entry)
XX      24-APR-2003 (first entry)
DE      Corn ear-derived polynucleotide (cpd) #463.
XX      Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
XX      structural gene; functional gene; regulatory gene;
XX      corn ear-specific profile; gene transcription; gene expression;
XX      hybrid plant; desirable trait expression; plant breeding program;
XX      inheritance; desirable characteristics; growth; development;
XX      disease resistance; environmental adaptability; quality; yield;
XX      multigene trait; plant; gene; ss.
XX      Zea mays.
OS      Zea mays.
XX      US6476212-B1.
XX      05-NOV-2002.
XX      14-MAY-1999; 99US-00313294.
XX      26-MAY-1998; 98US-0086722P.
XX      (INCY-) INCYTE GENOMICS INC.
XX      (INCY-) INCYTE GENOMICS INC.

```


CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

Query Match 2.1%; Score 68.6; DB 5; Length 556;
 Best Local Similarity 48.3%; Pred. No. 2.8e-06;
 Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

265 CTCGGTGGTCTGCTGTAAGTTTCTGCTTACAGCCCTGCGATCGCCCTCAGCGATGCA 324
 486 CCCCCCTCCCTTTTAAATATTTTATTTTCTCCCTCCCTCCCTCCATATA 427
 325 TTCGTCGCTGATTTCTTTTCTTCCGCTGGAATAATGCGCTATGTTCTGATTTGG 384
 426 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 367
 385 AAGTTTCTGCTGCTGTAAGTTTCTGCTTACAGCCCTGCGATCGCCCTCAGCGATGCA 444
 366 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 307
 445 CTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504
 306 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 247
 505 GATTTGACGCTGGAATAATCTTCGCTTTTCTGCTTTTCTGCTTATTAATGCGATTTGAT 564
 246 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 187
 565 CTACCTTTTCTGCTTGTGATTTGTTTGTGAGCCTAAGCGTGTGCTGCTGCTTATACT 624
 186 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
 625 TCACGCTTCACTGCTGATTTTGAATTTT 653
 126 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 98

RESULT 8

ABV42105/c
 ID ABV42105 standard; cDNA; 556 BP.

AC ABV42105;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 42096.

DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001MO-US005171.

PR 17-FEB-2000; 2000US-018319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PA Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

DR Novel isolated nucleic acid molecule associated with cancerous state of
 XX prostate cells and correlating with presence of prostate cancer; useful
 PT for detecting presence of prostate cancer; stage of prostate cancer.
 XX
 SS Claim 1; Page 8441; 11750p; English.

XX The invention relates to an isolated nucleic acid molecule (i) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (i) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX

SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

Query Match 2.1%; Score 68.6; DB 5; Length 556;
 Best Local Similarity 48.3%; Pred. No. 2.8e-06;
 Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

265 CTCGGTGGTCTGCTGTAAGTTTCTGCTTACAGCCCTGCGATCGCCCTCAGCGATGCA 324
 486 CCCCCCTCCCTTTTAAATATTTTATTTTCTCCCTCCCTCCCTCCATATA 427
 325 TTCGTCGCTGATTTCTTTTCTTCCGCTGGAATAATGCGCTATGTTCTGATTTGG 384
 426 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 367
 385 AAGTTTCTGCTGCTGTAAGTTTCTGCTTACAGCCCTGCGATCGCCCTCAGCGATGCA 444
 366 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 307
 445 CTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504
 306 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 247
 505 GATTTGACGCTGGAATAATCTTCGCTTTTCTGCTTTTCTGCTTATTAATGCGATTTGAT 564
 246 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 187
 565 CTACCTTTTCTGCTTGTGATTTGTTTGTGAGCCTAAGCGTGTGCTGCTGCTTATACT 624
 186 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
 625 TCACGCTTCACTGCTGATTTTGAATTTT 653
 126 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 98

RESULT 9

ABV43601/c
 ID ABV43601 standard; cDNA; 556 BP.

AC ABV43601;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 43592.

DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

KW Nijmegen breakage syndrome; Werner syndrome; immunodeficiency;
 KM trichothiodystrophy; Fanconi's anaemia; solid tumour; cancer; ds.
 XX
 OS Unidentified.
 XX
 PN WO200181622-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 06-APR-2001; 2001WO-EP003972.
 XX
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-034446/04.
 XX
 PT New nucleic acid derived from genes associated with DNA repair, useful
 PT for diagnosis, e.g. of ataxia telangiectasia, by determination of
 PT cytosine methylation.
 XX
 PS Claim 1; SEQ ID NO i22; 25pp + Sequence Listing; English.
 CC The invention relates to nucleic acids containing a sequence of at least
 CC 18 nucleotides of chemically treated DNA of genes associated with DNA
 CC repair, and their complements. The invention also relates to nucleic
 CC acids comprising at least 18 base pairs of the chemically pretreated DNA
 CC of genes associated with DNA repair selected from PMS2L1, PMS2L2,
 CC PMS2L3, PMS2, L4, PMS2L5, MGMT, MSH2, NUDT1, TDG, INPPL1,
 CC RFC4, DIRTLL, FANCB, or XRCC8. Nucleic acids of the invention and related
 CC oligomers, are useful for diagnosis of diseases associated with gene
 CC repair, specifically ataxia telangiectasia, aging, Bloom's syndrome,
 CC Cockayne syndrome, Nijmegen breakage syndrome or Werner syndrome.
 CC immunodeficiency, trichothiodystrophy, Fanconi's anaemia, solid tumours
 CC and cancer, particularly by determining status of cytosine methylation
 CC and/or by detecting single-nucleotide polymorphisms. Determination of
 CC individual methylation patterns may allow development of individualised
 CC therapies. The sequences given in records AB92192-AB92335 represent
 CC chemically pre-treated DNA fragments from genes associated with DNA
 CC repair, and their complements. Note: The sequence data for this patent is
 CC not represented in the specification, but is based on sequence
 CC information supplied by the European Patent Office
 CC
 SQ Sequence 8079 BP; 1992 A; 212 C; 1962 G; 3913 T; 0 U; 0 Other;
 Query Match 2.0%; Score 66.4; DB 6; Length 8079;
 Best Local Similarity 48.2%; Pred. No. 3.2e-05;
 Matches 187; Conservative 0; Mismatches 201; Indels 0; Gaps 0;
 QY 333 TCGATTTCTTTTCTTCGCGTGAAGAAATGCGCTTAAGTTCGATTCGAAGGTTT 392
 DB 5327 TTGTTTATGTTTATATATATAGTGAAGCGTATATGTTATGTTATTTGTTT 5386
 QY 393 TGTGCTAGGGTACTTTTTCCTTATTTTAAATTTTAAAGTTCTAGTAACGATTCGCTC 452
 DB 5387 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5446
 QY 453 TTAAGTCTTTTGTTCATTTTGTGTCCTTACCGTTTACGCTGATCGAGATTGTA 512
 DB 5447 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5506
 QY 513 CTGTGAAGAAATCCTTGTTTGTGTTTGTTCATTAATTAATCGATGATCTTACCTT 572
 DB 5507 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5566
 QY 573 TGTGCTTAGATGTTTGTGTTTGTGAGCCATGCGTGTGGCTGTATAACTTCACGCTC 632
 DB 5567 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5626

QY 633 ATGTGTGATTTTGAGATTTTGGTAGTACGTGGGTTTCTTGTGCGTATAGCTTGA 692
 DB 5627 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5686
 QY 693 AAAATGAGTTCTAGGCGTGTGCAATGT 720
 DB 5687 GTGTAGTGTATGTTATGTTATTTGT 5714
 RESULT 13
 AAX33181/C
 ID AAX33181 standard; DNA; 6644 BP.
 XX
 AC AAX33181;
 XX
 DT 25-JUN-1999 (first entry)
 XX
 DE Base sequence of the plasmid pRX-ires-bar.
 XX
 KM Compo virus; bsr; viral vector; expression; apoptosis; resistance; crna;
 KM bcl-2; bcl-x1; FLIP; survivin; IAP; ILP; adenovirus; cancer;
 KM autoimmune disease; graft rejection reaction; inflammation;
 KM inflammatory disease; ss.
 XX
 OS Synthetic.
 OS Compo virus.
 XX
 PN WO9913073-A2.
 XX
 PD 18-MAR-1999.
 XX
 PF 07-SEP-1998; 98WO-UP004010.
 XX
 PR 08-SEP-1997; 97JP-00259235.
 XX
 PA (RPRG-) RPR GENCELL ASIA PACIFIC INC.
 XX
 PI Hamada H;
 XX
 DR WPI; 1999-243728/20.
 XX
 PT New apoptosis-resistant virus-sensitive cell.
 XX
 PS Example 1; Page 38-41; 51pp; English.
 XX
 CC The present invention describes an apoptosis-resistant virus-sensitive
 CC cell line into which an apoptosis resistance gene has been introduced.
 CC The recombinant viruses generated are capable of expressing apoptosis-
 CC associated genes. These can then be used in a variety of diseases for
 CC which the induction of apoptosis by gene transfer, or where the
 CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
 CC are useful as vectors for gene therapy which can be applied to cancer
 CC therapy for destroying cancer cells selectively, the treatment of
 CC autoimmune diseases and graft rejection reaction, and apoptosis induction
 CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
 CC encountered the problem where if an adenovirus vector capable of
 CC expressing an apoptosis-associated gene is introduced into animal cells,
 CC the cells producing the virus will be destroyed because the period of
 CC time required to induce cell death by apoptosis is shorter than that
 CC required to replicate and produce the virus, resulting in failure to
 CC obtain a recombinant virus having the integrated apoptosis-associated
 CC gene. In this invention an apoptosis-resistant 293 cell line (having an
 CC apoptosis resistant gene introduced) is established and overcomes the
 CC problem. The present sequence represents the base sequence of the plasmid
 CC pRX-ires-bar, which contains the compo virus bsr gene, and is used in an
 CC example from the present invention
 CC
 SQ Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 U; 0 Other;
 Query Match 2.0%; Score 66; DB 2; Length 6644;
 Best Local Similarity 45.4%; Pred. No. 3.6e-05;
 Matches 237; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 14:30:53 ; Search time 8548.15 Seconds
(without alignments)
16605.743 Million cell updates/sec

Title: US-10-030-829-1
Perfect score: 3275
Sequence: 1 gacaaacaacaaatctaa.....caaatatagggtacaat 3275

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank1:
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sfg:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3275	100.0	3275	6	AX078760
2	3275	100.0	81265	8	AB025633
3	2254	68.8	2254	8	AF239719
4	1651.2	50.4	2162	8	BT002944
5	1493	45.6	1909	8	BT004380
6	1462	44.6	1878	6	AX078761
7	643.6	19.7	650	8	ATMS28171
8	194	5.9	1556	8	AK064217
9	189.8	5.8	2181	8	AK064995
10	181.8	5.6	94911	8	ATT14E10
11	174	5.3	91053	8	CNS08CNO
12	174	5.3	112231	8	CNS07YOL
13	155.8	4.8	2570	8	AK100659
14	120.8	3.7	1564	8	AF542974
15	117	3.6	513	8	AF469493
16	112.8	3.4	1214	8	AK061975
17	85.6	2.6	7218	6	166494
18	83.4	2.5	184535	6	CNS05TCT
19	81.4	2.5	34980	6	AX344566
20	74.4	2.3	272	6	AR244762
21	72.8	2.2	1407	8	AJ592026
22	72.2	2.2	274	6	AR245104
23	71	2.2	60565	2	AC023852
24	70	2.1	1453	8	AJ591978
25	70	2.1	66993	2	AC138074
26	69.6	2.1	1434	8	AJ592058
27	69.2	2.1	3364	14	TSP418778
28	69	2.1	88932	2	AC022680
29	68.8	2.1	6668	6	AX346599
30	68.8	2.1	81120	2	AC022851
31	68.6	2.1	7218	6	166494
32	68.4	2.1	810	8	AJ592180
33	68.2	2.1	56857	2	AC021917
34	68.2	2.1	194180	2	AC069140
35	68	2.1	1184	8	AJ591982
36	67.6	2.1	151368	2	AC142664
37	67.4	2.1	156550	2	AC015830
38	67.4	2.1	196853	9	CNS01DUD
39	67.4	2.1	273807	2	AC025421
40	67	2.0	1228	8	AJ592059
41	67	2.0	127709	2	CNS05TCT
42	67	2.0	202083	3	AC023833
43	66.8	2.0	9539	6	AX277889
44	66.8	2.0	9539	6	AX323566
45	66.8	2.0	70389	2	AC135853

ALIGNMENTS

RESULT 1
AX078760 3275 bp DNA linear PAT 22-FEB-2001
LOCUS Sequence 1 from Patent WO0105951.
DEFINITION AX078760
ACCESSION AX078760.1 GI:1158379
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS
TITLE
1 Beclin,C., Elmayan,T. and Vaucheret,H.
Novel gsg3 plant gene and use thereof

Pred. No. is the number of results predicted by chance to have a

QY 1921 GCTGGGGGTCAGAGCGCAGTATGTTTTCTGAGGGTTCGCCAATCTGATGGCTTCTT 1980
 DB 1921 GCTGGGGGTCAGAGCGCAGTATGTTTTCTGAGGGTTCGCCAATCTGATGGCTTCTT 1980
 QY 1981 GCAGAGAGCAAGATCTGGACATATTCATCAACACTCTCAAGGTTCTCTCCCAAGAA 2040
 DB 1981 GCAGAGAGCAAGATCTGGACATATTCATCAACACTCTCAAGGTTCTCTCCCAAGAA 2040
 QY 2041 AATTTGATATATGCTTTTATGTTTGTGATGAAATTAAGTTTGTGTCGCTGTTAA 2100
 DB 2041 AATTTGATATATGCTTTTATGTTTGTGATGAAATTAAGTTTGTGTCGCTGTTAA 2100
 QY 2101 TGCATCTGTTATATATATCTATGATTCATTAAGCAAAACAAGGCTGAATTCAGTTG 2160
 DB 2101 TGCATCTGTTATATATATCTATGATTCATTAAGCAAAACAAGGCTGAATTCAGTTG 2160
 QY 2161 AATCATACCAAGATGTTGTTAAAGAGCTGAGGCAATCTCTAGAGCAATCAAGCAG 2220
 DB 2161 AATCATACCAAGATGTTGTTAAAGAGCTGAGGCAATCTCTAGAGCAATCAAGCAG 2220
 QY 2221 CTGAATCTATTAAGAACAGGCTCAAAAACGAACAACGCAAGGTCGTAGAGAA 2280
 DB 2221 CTGAATCTATTAAGAACAGGCTCTCAAAAACGAACAACGCAAGGTCGTAGAGAA 2280
 QY 2281 TCTCTGAAATTAATGAGCGAGAGCTCGTAGAATCTGAGAGATTAATCGATCGTGA 2340
 DB 2281 TCTCTGAAATTAATGAGCGAGAGCTCGTAGAATCTGAGAGATTAATCGATCGTGA 2340
 QY 2341 CAGAGACTAAGATGAGCATGAACAGAACAGGAAAGAGTATGATTTTCTAGAAAT 2400
 DB 2341 CAGAGAACTAAGATGAGCATGAACAGAACAGGAAAGAGTATGATTTTCTAGAAAT 2400
 QY 2401 CACAACTTGATGATTTGATTAACCACTGATTCACATTTTGTATATATGTCACAA 2460
 DB 2401 CACAACTTGATGATTTGATTAACCACTGATTCACATTTTGTATATATGTCACAA 2460
 QY 2461 AAAACCTGTGTGTTGGAAGATGATGACACAGAGGTTTTCATGATTCATCAAA 2520
 DB 2461 AAAACCTGTGTGTTGGAAGATGATGACACAGAGGTTTTCATGATTCATCAAA 2520
 QY 2521 CAGATTCATGAAGAGAGAGCGAAAGAGAGATTTGAGATGTTGACAGCAGAGAA 2580
 DB 2521 CAGATTCATGAAGAGAGAGCGAAAGAGAGATTTGAGATGTTGACAGCAGAGAA 2580
 QY 2581 CGTGCCAAAGTGTGTGGCCAGAGCAGAGCAATTAATCCCTTAGCAATGACGATTC 2640
 DB 2581 CGTGCCAAAGTGTGTGGCCAGAGCAGAGCAATTAATCCCTTAGCAATGACGATTC 2640
 QY 2641 CGAAAGAGGTATATGTAATCAATATATCCCTGCGTTTGTGTTTCAAACTTA 2700
 DB 2641 CGAAAGAGGTATATGTAATCAATATATCCCTGCGTTTGTGTTTCAAACTTA 2700
 QY 2701 AGAGTACTGAATATATCCGTTTGTGATTCCTTCGAGAGCTGAGAGAGTTCACCTTC 2760
 DB 2701 AGAGTACTGAATATATCCGTTTGTGATTCCTTCGAGAGCTGAGAGAGTTCACCTTC 2760
 QY 2761 ATCGAGTTTCAAGAGAGAGATGAGAGGTTTGTGAGAGAGGAGATGCTGATTA 2820
 DB 2761 ATCGAGTTTCAAGAGAGAGATGAGAGGTTTGTGAGAGAGGAGATGCTGATTA 2820
 QY 2821 GATCAAGAGAGAGATGAGAGAGATGAGAGAGGATCAAGAGAGATTAATTTGATCTG 2880
 DB 2821 GATCAAGAGAGAGATGAGAGAGATGAGAGAGGATCAAGAGAGATTAATTTGATCTG 2880
 QY 2881 GAGAAAGAAATTTGATGAGGCTTTGAAACAGCTCATGTAAGAGATGCTTCAATGAA 2940
 DB 2881 GAGAAAGAAATTTGATGAGGCTTTGAAACAGCTCATGTAAGAGATGCTTCAATGAA 2940
 QY 2941 GATGATGAGAGAGAGATGATGATGACAGAGAGATGATGATGATGATGATGATG 3000
 DB 2941 GATGATGAGAGAGAGATGATGATGACAGAGAGATGATGATGATGATGATGATG 3000
 QY 3001 GTATGTCCGAAAGTAGAGATCTGAGAGATCTCAATTAATTAATGACAAATCTAAGGA 3060

DB 3001 GTATGTCCGAAAGTAGAGATCTGAGAGATCTCAATTAATTAATGACAAATCTAAGGA 3060
 QY 3061 GATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3120
 DB 3061 GATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3120
 QY 3121 GTGACTAAACCAAGTTTCTAGTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 3180
 DB 3121 GTGACTAAACCAAGTTTCTAGTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 3180
 QY 3181 TTAGACATTTTACCAAGCTCAGAGTGAATCAAGATGCGCAATCAATATGATTTT 3240
 DB 3181 TTAGACATTTTACCAAGCTCAGAGTGAATCAAGATGCGCAATCAATATGATTTT 3240
 QY 3241 AGAATTTTATCTCAAAATTAATGAGTACAAAT 3275
 DB 3241 AGAATTTTATCTCAAAATTAATGAGTACAAAT 3275

RESULT 2
 AB025633
 LOCUS AB025633 81365 bp DNA linear PLN 27-DEC-2000
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:M0M1.
 ACCESSION AB025633 BA000015
 VERSION AB025633.2 GI:10178221
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (sites)

REFERENCE
 AUTHORS Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., Kotani, H. and Tabata, S.
 TITLE Structural analysis of Arabidopsis thaliana chromosome 5. X.
 Sequence features of the regions of 3,076,755 bp covered by sixty
 P1 and TAC clones
 JOURNAL DNA Res. 7 (1), 31-63 (2000)
 MEDLINE 20181125
 PUBMED 10718197

REFERENCE
 AUTHORS Nakamura, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-1999) Yasukazu Nakamura, Kazusa DNA Research
 Institute, Department of Plant Gene Research, 1532-3, Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
 Tel:81-438-52-3935, Fax:81-438-52-3934)
 On Sep 15, 2000 this sequence version replaced gi:4589439.
 Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=M0M1

COMMENT

Genes with similarity to proteins in the databases are described in
 'product' or 'note' qualifiers. Genes that have no significant
 protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Graal
 (Informatics Group, Oak Ridge National Laboratory,
 http://compbio.ornl.gov/Graal-1.3/),
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University,
 http://gremli1.zool.iastate.edu/cgi-bin/dsp.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-SE/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is K19M13 and the 3' clone is MRO11.

FEATURES

source
 1..81365
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"

[illegible]


```

CDS      complement(join(29439..29576,29661..29885,29985..30268,
30341..30493,30579..30722,30956..31484))
        /note="gene_id:MGM1.15"

```


Db 61 GAGGTGAACAGTGGTTCAAGGTTGGCAGGAGCAGACTGGCTTCTTCAACAATGAT 120
 Qy 816 GAGAGAGTGGGAGGTCATTTTCCAGAGAAACAACAACAACAGAGAAACATTTCTGA 875
 Db 121 GAGGAGAGTGGAGGTCAATTTCCAGAGAAACAACAACAACAGAGAAACATTTCTGA 180
 Qy 876 AAAACTTGGGTTCTCAGAAATTCGATCTCTAGAGCTTGGGGTGTGACGACAGAGG 935
 Db 181 AAAACTTGGGTTCTCAGAAATTCGATCTCTAGAGCTTGGGGTGTGACGACAGAGG 240
 Qy 936 AGAGTAGCAACGTAATCTGGAGAGAGAAACAATGTAACCGAGAGATACCGCAATGGT 995
 Db 241 AGAGTAGCAACGTAATCTGGAGAGAGAAACAATGTAACCGAGAGATACCGCAATGGT 300
 Qy 996 CCGGGCATTCAGACTAACATATCTGGTCCGGGACGAGCTTGGAGAGAAAGTATGTAAC 1055
 Db 301 CCGGGCATTCAGACTAACATATCTGGTCCGGGACGAGCTTGGAGAGAAAGTATGTAAC 360
 Qy 1056 AACTTGTGGCACCCACCTGTATCTGGCCCTCTTGGAGAGAGATGGAATGGACAG 1115
 Db 361 AACTTGTGGCACCCACCTGTATCTGGCCCTCTTGGAGAGAGATGGAATGGACAG 420
 Qy 1116 GCAAGAGAGGTTCTGCTCAGCAACAGCTGTGACAGAGTTTCTGACGTGAGATGAT 1175
 Db 421 GCAAGAGAGGTTCTGCTCAGCAACAGCTGTGACAGAGTTTCTGACGTGAGATGAT 480
 Qy 1176 GTGAGTAATCTTCTGAGAGAGAGATGATCCGATGCTTTGATGATATCTGATGACAC 1235
 Db 481 GTGAGTAATCTTCTGAGAGAGAGATGATCCGATGCTTTGATGATATCTGATGACAC 540
 Qy 1236 CTTCGAAAGTATGATATGATGATCGAGATGATCAAAAAGAGCCATGATCAAGAAAGCAG 1295
 Db 541 CTTCGAAAGTATGATATGATGATCGAGATGATCAAAAAGAGCCATGATCAAGAAAGCAG 600
 Qy 1296 AATAAGTGTTCAAAAAGTTCTTTGGCAGCTTGGATAGCTTGTGATCGACAGATTAAT 1355
 Db 601 AATAAGTGTTCAAAAAGTTCTTTGGCAGCTTGGATAGCTTGTGATCGACAGATTAAT 660
 Qy 1356 GAACCAACAAGGACAGTGGCATTTCCAGCTTGTCAACAACGAGCTGTGCTCATCGATTGG 1415
 Db 661 GAACCAACAAGGACAGTGGCATTTCCAGCTTGTCAACAACGAGCTGTGCTCATCGATTGG 720
 Qy 1416 TTTAACCTGACACCTCTACTAGCTCATGCGAGAGCAAAAAGAGCTGAGGAGTTAAGCTC 1475
 Db 721 TTTAACCTGACACCTCTACTAGCTCATGCGAGAGCAAAAAGAGCTGAGGAGTTAAGCTC 780
 Qy 1476 CATAGAGAAATGGCTGAGATTTTGAAGAAAGATCTACAGATGAGAGCGCATCTGCATT 1535
 Db 781 CATAGAGAAATGGCTGAGATTTTGAAGAAAGATCTACAGATGAGAGCGCATCTGCATT 840
 Qy 1536 CCTTGTGTGAGATTTTATGAGGACGTGGAAGGTTTGGTGAAGATTAAGAAAGTTATGAA 1595
 Db 841 CCTTGTGTGAGATTTTATGAGGACGTGGAAGGTTTGGTGAAGATTAAGAAAGTTATGAA 900
 Qy 1596 ATTGTCGTGCTCCATAGTCATCATGAAATCTAGACTGGAATAAGACGATACGAT 1655
 Db 901 ATTGTCGTGCTCCATAGTCATCATGAAATCTAGACTGGAATAAGACGATACGAT 960
 Qy 1656 AAGGTGGAATTTCTTGTCTTTTACTTTTACTTTTCTTTGACATTTCTACTGATCTTA 1715
 Db 961 AAGGTGGAATTTCTTGTCTTTTACTTTTACTTTTCTTTGACATTTCTACTGATCTTA 1020
 Qy 1716 GAATGTTAATTTGATGAGTGGCTGGCATGGGCAACAAGAGCTGTGGAATCTTGCACAA 1775
 Db 1021 GAATGTTAATTTGATGAGTGGCTGGCATGGGCAACAAGAGCTGTGGAATCTTGCACAA 1080
 Qy 1776 GTATGAGGCTTTTGAAGACGACATCTTATGTTCCACAGGGCATCGTGGAGTGAATGT 1835
 Db 1081 GTATGAGGCTTTTGAAGACGACATCTTATGTTCCACAGGGCATCGTGGAGTGAATGT 1140
 Qy 1836 TCTGATGTTTGAAGAGTGCATGCTGCTATTTTGAAGGCGCAACGCTCCACCGGAGTT 1895
 Db 1895 TCTGATGTTTGAAGAGTGCATGCTGCTATTTTGAAGGCGCAACGCTCCACCGGAGTT 1895

Db 1141 TCTGATGTTTGAAGACAGTGCACCTGCTATTTTGAAGGCGCAACGCTCCACCGGAGTT 1200
 Qy 1896 AGCTGATGAGGTTAGATAGAAATTTGCCCTGGGGTCAAGAGCGCAGATGTTTCTGAGAG 1955
 Db 1201 AGCTGATGAGGTTAGATAGAAATTTGCCCTGGGGTCAAGAGCGCAGATGTTTCTGAGAG 1260
 Qy 1956 TGTTCGCCAATCTGATGAGCTTCTTCCAAAGAGAGATCTGACATATTCATCAACA 2015
 Db 1261 TGTTCGCCAATCTGATGAGCTTCTTCCAAAGAGAGATCTGACATATTCATCAACA 1320
 Qy 2016 CTCTCAAGGTTCTCTCCCAAGAAATTTGATATATGCTTTTATGTTTGTCAATGGAAT 2075
 Db 1321 CTCTCAAGGTTCTCTCCCAAGAAATTTGATATATGCTTTTATGTTTGTCAATGGAAT 1380
 Qy 2076 TTTAAGTTTGTGTCCTGTTAATGATCTGTTATGATATCTATGATTTCAATTAAG 2135
 Db 1381 TTTAAGTTTGTGTCCTGTTAATGATCTGTTATGATATCTATGATTTCAATTAAG 1440
 Qy 2136 CAAAACAAGGCTGAATTCGAGTTGAAATCATACCAAGAGATGTTGTAAGGAGCTGAG 2195
 Db 1441 CAAAACAAGGCTGAATTCGAGTTGAAATCATACCAAGAGATGTTGTAAGGAGCTGAG 1500
 Qy 2196 GCAGATCTCTGAGACCAATCAGCAGCTGAATCTTTAAGAACAACTCTCAAAACAGAA 2255
 Db 1501 GCAGATCTCTGAGACCAATCAGCAGCTGAATCTTTAAGAACAACTCTCAAAACAGAA 1560
 Qy 2256 CAGACCGCAGAGGCTTGAAGAAATCTGTGAATTTATGACGGAAGCTGCTGAAGAC 2315
 Db 1561 CAGACCGCAGAGGCTTGAAGAAATCTGTGAATTTATGACGGAAGCTGCTGAAGAC 1620
 Qy 2316 TGCAGAGATTAATCGGATCGTGAACAGAGAACTAAGATGACAGTGAACAGAAACAGGGA 2375
 Db 1621 TGCAGAGATTAATCGGATCGTGAACAGAGAACTAAGATGACAGTGAACAGAAACAGGGA 1680
 Qy 2376 AAGGATGATTTTCTTGAAGAAATCACAAATTTGATATTTACTGATTTCA 2435
 Db 1681 AAGGATGATTTTCTTGAAGAAATCACAAATTTGATATTTACTGATTTCA 1740
 Qy 2436 CATTTTGAATTAATTTGTCACAAAAAACCTGTGTGTTTGAAGATGACACAGCA 2495
 Db 1741 CATTTTGAATTAATTTGTCACAAAAAACCTGTGTGTTTGAAGATGACACAGCA 1800
 Qy 2496 CAGTTTTTTCATGATTCATCAACAAGATCCATGAAAGAGAGACGCAAGAGAGAA 2555
 Db 1801 CAGTTTTTTCATGATTCATCAACAAGATCCATGAAAGAGAGACGCAAGAGAGAA 1860
 Qy 2556 TTTGAGATGTTGACAGCAGACAGAAAGTGCAGAGGTTTGGCCAGCAGCAGCAACAT 2615
 Db 1861 TTTGAGATGTTGACAGCAGACAGAAAGTGCAGAGGTTTGGCCAGCAGCAGCAACAT 1920
 Qy 2616 TTAATCCCTCTAGCATGACATTTGCCGAAAGAGGTTATGTTCTTAACATTAATCCCT 2675
 Db 1921 TTAATCCCTCTAGCATGACATTTGCCGAAAGAGGTTATGTTCTTAACATTAATCCCT 1980
 Qy 2676 CTGGCGTTTTTGTTTTCAAACTTAAGATTAACGTAATTTCCGAGTTTGTGATCTTTG 2735
 Db 1981 CTGGCGTTTTTGTTTTCAAACTTAAGATTAACGTAATTTCCGAGTTTGTGATCTTTG 2040
 Qy 2736 CAGAGCTGAGAAAGTGCAGAGCTTCATGAGTTTCAAGAGAAAGATGAGAGGTTGT 2795
 Db 2041 CAGAGCTGAGAAAGTGCAGAGCTTCATGAGTTTCAAGAGAAAGATGAGAGGTTGT 2100
 Qy 2796 GGAAGAGAGGAGATGCTGATTAAGATCAAGAGAAAGAGATGAGAAACATGAAGAAG 2855
 Db 2101 GGAAGAGAGGAGATGCTGATTAAGATCAAGAGAAAGAGATGAGAAACATGAAGAAG 2160
 Qy 2856 GCATCAGAGAGATATTTGATCTGAGAGAAAGATTTTATGAGGCTTTGGAACAGCTCAT 2915
 Db 2161 GCATCAGAGAGATATTTGATCTGAGAGAAAGATTTTATGAGGCTTTGGAACAGCTCAT 2220
 Qy 2916 GTACAGCATGGCTTCACAATGAAGATGATGA 2949
 Db 2221 GTACAGCATGGCTTCACAATGAAGATGATGA 2254

RESULT 4
LOCUS BT002944
DEFINITION Arabidopsis thaliana clone RAF14-93-K05 (R20243) unknown protein (At5g23570) mRNA, complete cds.
ACCESSION BT002944
VERSION BT002944.1
KEYWORDS GI:27754622
SOURCE FLI CDNA.
ORGANISM Arabidopsis thaliana (thale cress)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2162)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 2162)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: RIKEN Arabidopsis Full-length cDNA): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEBC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.
Yamada,K. (SSP/PGEBC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEBC) contributed equally to this work as PIs.
Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.
FEATURES
source
1..2162
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAF14-93-K05 (R20243)"
/ecotype="Columbia"
/note="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert."
gene
1..2162
/gene="At5g23570"
5'UTR
1..91
/gene="At5g23570"
CDS
92..1969
/gene="At5g23570"
/codon_start=1

ORIGIN
Query Match 50.4%; Score 1651.2; DB 8; Length 2162;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 2069; Conservative 0; Mismatches 3; Indels 376; Gaps 4;
misc_difference 2146
/gene="At5g23570"
/note="not present in genomic sequence"
3'UTR
1970..2162
/gene="At5g23570"
misc_difference 2146
/gene="At5g23570"
/note="not present in genomic sequence"
QY 687 GTTGAATAAAGAGTCTAGAGGCTGTCATGCTTAAGAAAGAAAGCTTCAGGGTGGT 746
DB 83 GTTGAATAAAGAGTCTAGAGGCTGTCATGCTTAAGAAAGAAAGCTTCAGGGTGGT 142
QY 747 TATAGGCTGAGGTGGAACAGTGTTCAGAGTTTGCGAGGAGCAGACTGCTTCTTCA 806
DB 143 TATAGGCTGAGGTGGAACAGTGTTCAGAGTTTGCGAGGAGCAGACTGCTTCTTCA 202
QY 807 CAGATGATGAGAGAGTGGAGGTCTTTCCAAAGAAACAAACCAAGCAAGAAC 866
DB 203 CAGATGATGAGAGAGTGGAGGTCTTTCCAAAGAAACAAACCAAGCAAGAAC 262
QY 867 ACTTCGGAAGAACTTGGGTTTCTCAGAAATCGAATCCTCTAGAGCTTGGGGTGGT 926
DB 263 ACTTCGGAAGAACTTGGGTTTCTCAGAAATCGAATCCTCTAGAGCTTGGGGTGGT 322
QY 927 CAGCAAGGAGAGAGTGAACATGATCTGCGAGAGAAACAATGTTCCGGAGAGGTAA 986
DB 323 CAGCAAGGAGAGAGTGAACATGATCTGCGAGAGAAACAATGTTCCGGAGAGGTAA 382
QY 987 GGCATGCTCGGGGATTCAGACTAATCTGCTGCGGGGACGAGCGTTGAGCAGAAAG 1046
DB 383 GGCATGCTCGGGGATTCAGACTAATCTGCTGCGGGGACGAGCGTTGAGCAGAAAG 442
QY 1047 TATGATTAACAATTTGTGCAACCCCACTGATCTGCGCTCTTTGGAAGAGAGATGG 1106
DB 443 TATGATTAACAATTTGTGCAACCCCACTGATCTGCGCTCTTTGGAAGAGAGATGG 502
QY 1107 AATTGCGAGGCAAGAGAGCTTCTGTCAGCAACAGCTGTGACAGAGTTCTGACG 1166
DB 503 AATTGCGAGGCAAGAGAGCTTCTGTCAGCAACAGCTGTGACAGAGTTCTGACG 562
QY 1167 GAGGATGATGGGATTAATGCTTCTGAGAAAGAAATGATTCGATGCTTGGATGAT 1226
DB 563 GAGGATGATGGGATTAATGCTTCTGAGAAAGAAATGATTCGATGCTTGGATGAT 622
QY 1227 GATGAGACCTTGCAAGTATGATTTGATCTGAGTGTGATCAAAAGCCATGATTA 1286
DB 623 GATGAGACCTTGCAAGTATGATTTGATCTGAGTGTGATCAAAAGCCATGATTA 682
QY 1287 CGAAAGCAGAAATAGGCTTCAAAAGCTTTGGAGCTTGGATGCTTGTGATGAG 1346
DB 683 CGAAAGCAGAAATAGGCTTCAAAAGCTTTGGAGCTTGGATGCTTGTGATGAG 742
QY 1347 CAGATAAATGAACCAAGAGGAGTGGATGCTTGTGATGAGACGAGCTGTGTC 1406

Db 743 CAGATAAATGACCAAGAGGAGTGGCATTGTCCAGCTTGTCAAGACGACCTGTGCC 802
 QY 1407 ATGATTTGGTATTAACCTGCACCCCTCTACTAGCTCATGCGAGGACAAAGAGGCTAGGCCGA 1466
 Db 803 ATGATTTGGTATTAACCTGCACCCCTCTACTAGCTCATGCGAGGACAAAGAGGCTAGGCCGA 862
 QY 1467 GTTAACTCCATAGAGAAATTTGGCTGAAGTTTAAAGAAAGAGTCTACAGATGAGAGGCGCA 1526
 Db 863 GTTAACTCCATAGAGAAATTTGGCTGAAGTTTAAAGAAAGAGTCTACAGATGAGAGGCGCA 922
 QY 1527 TCTGTCTCTTCTGTGTGTGAGATTTATGGGCACTGGAAGGTTTGGGTGAGATGAAG 1586
 Db 923 TCTGTCTCTTCTGTGTGTGAGATTTATGGGCACTGGAAGGTTTGGGTGAGATGAAG 982
 QY 1587 GATTATGAATTTGTGGGCTCCAAATGGTCAATCATATGAATCTAGCTGATGAAGGAC 1646
 Db 983 GATTATGAATTTGTGGGCTCCAAATGGTCAATCATATGAATCTAGCTGATGAAGGAC 1042
 QY 1647 GATTAAGATTAAGGTGAATTTCTGTCTTTTACTTTTAAATTTTCTTGTCAATCTTA 1706
 Db 1043 GATTAAGATTA----- 1052
 QY 1707 CTGATCTTGAATTTTACTATTGATGTGGCTGGGCAATGGGCAACAAAGCTGTGAAATA 1766
 Db 1053 -----AGTGGCTGGGCAATGGGCAACAAAGCTGTGAAATA 1089
 QY 1767 CTTCGACAGTATGAGGCTCTTAGACAGCCCAATTCCTATAGTCCACAGGCGCATGTGG 1826
 Db 1090 CTTCGACAGTATGAGGCTCTTAGACAGCCCAATTCCTATAGTCCACAGGCGCATGTGG 1149
 QY 1827 GATGAGTGTCTGATTTTGAAGAGCAGTGGCCATGCTATTTGAGAGCCGAACGCTCCCA 1886
 Db 1150 GATGAGTGTCTGATTTTGAAGAGCAGTGGCCATGCTATTTGAGAGCCGAACGCTCCCA 1209
 QY 1887 CCGGAGTTAAGCTGAGATGGGGTTAATGAATTTGCTGGGGTCAAGAGCGCATATGTT 1946
 Db 1210 CCGGAGTTAAGCTGAGATGGGGTTAATGAATTTGCTGGGGTCAAGAGCGCATATGTT 1269
 QY 1947 TTCTGAGAGTGTGGGCAACTGATGAGGCTTCTTGCAAGAGCAAGATCTGACATATT 2006
 Db 1270 TTCTGAGAGTGTGGGCAACTGATGAGGCTTCTTGCAAGAGCAAGATCTGACATATT 1329
 QY 2007 CAATCAACACTCTCAAGGTTCTCTCCCAAGAAATTTGATATATGCTTTAGTTTGT 2066
 Db 1330 CAATCAACACTCTCA----- 1344
 QY 2067 CATTGGAATTTAAAGTTTGTGTGCTCCGTTAATGCAATCTGTTATATATCTATGA 2126
 Db 1345 ----- 1344
 QY 2127 TTCATTAGGCAAAAGGCTGAATTTGAGTTGAATCATACAGAGATGGTTGTAAA 2186
 Db 1345 -----AGGCAAAAGGCTGAATTTGAGTTGAATCATACAGAGATGGTTGTAAA 1398
 QY 2187 GGAGCTGAGGAGATCTCTGAGGACATGACAGCTGAATCTATTTAAGAACAGCTCTC 2246
 Db 1399 GGAGCTGAGGAGATCTCTGAGGACATGACAGCTGAATCTATTTAAGAACAGCTCTC 1458
 QY 2247 AAAACAGAACAGACCGCAAGGCTCTTGAAGAAATCTTGAAGAAATTAAGAGCAAGCT 2306
 Db 1459 AAAACAGAACAGACCGCAAGGCTCTTGAAGAAATCTTGAAGAAATTAAGAGCAAGCT 1518
 QY 2307 GGGTAAAGCTGAGAGATTAATGAGATCTGAGAGACAGAACTAAGATGAGCATGAACA 2366
 Db 1519 GGGTAAAGCTGAGAGATTAATGAGATCTGAGAGACAGAACTAAGATGAGCATGAACA 1578
 QY 2367 GAAACAGGAAGGTATGATTTTCTTGAAGAAATCAAAACTGACATTTTGTATTACT 2426
 Db 1579 GAAACAGGAAG----- 1589
 QY 2427 ACTGATTCACATTTTGTATTAATTTGTCCAAACAAAACCTGTGTGTTGAAGATGA 2486
 Db 1590 -----AGATGGA 1596

QY 2487 TGCACACGACAGGTTTTTTCATGATTCATCAACAGATCCATGAAGAAAGAGACGCCAA 2546
 Db 1597 TGCACACGACAGGTTTTTTCATGATTCATCAACAGATCCATGAAGAAAGAGACGCCAA 1656
 QY 2547 GGAGGAGAAATTTGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2606
 Db 1657 GGAGGAGAAATTTGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1716
 QY 2607 GCAGAACATTAATCCCTCTAGCAATGACATTTGCCGAAGAGATATATGACTAATAC 2666
 Db 1717 GCAGAACATTAATCCCTCTAGCAATGACATTTGCCGAAG----- 1756
 QY 2667 ATATATCCCTGTGGCGTTTTTGTTTTCAAACTAAGATGATTAATTCGGTTTTG 2726
 Db 1757 ----- 1756
 QY 2727 ATCTTTGCGAGAGCTGAGAAAGTGTCAAGCTTCATCGAGTTTCAAGAGAAAGATGA 2786
 Db 1757 -----AGAGCTGAGAGAGTGTCAAGCTTCATCGAGTTTCAAGAGAAAGATGA 1806
 QY 2787 GGAGTTTGTGAAAGAGAGAGAGAGATGCTGTAATAAGATCAAGAGAGAGATGAGACAT 2846
 Db 1807 GGAGTTTGTGAAAGAGAGAGAGATGCTGTAATAAGATCAAGAGAGAGATGAGACAT 1866
 QY 2847 GAAAGAGAGAGATCAAGAGAGATTTTGAATCTGGAGAAAGATTTGATGAGGCTTTGA 2906
 Db 1867 GAAAGAGAGAGATCAAGAGAGATTTTGAATCTGGAGAAAGATTTGATGAGGCTTTGA 1926
 QY 2907 ACAGCTCATGTAACAGATGAGCTTCAACATGAAGATGATTAAGACAAAGTCTGTACA 2966
 Db 1927 ACAGCTCATGTAACAGATGAGCTTCAACATGAAGATGATTAAGACAAAGTCTGTACA 1986
 QY 2967 CAAGACAAACTAATGTTTCTTTGTTTGTCTTTGTTGATGTCCGAAGTGAAGATTTGAG 3026
 Db 1987 CAAGACAAACTAATGTTTCTTTGTTTGTCTTTGTTGATGTCCGAAGTGAAGATTTGAG 2046
 QY 3027 AGACTCCATTTAATACTAGCAACAATCTAAGAGATTAATGATTAATTTCCCTCAATT 3086
 Db 2047 AGACTCCATTTAATACTAGCAACAATCTAAGAGATTAATGATTAATTTCCCTCAATT 2106
 QY 3087 TTAGTAGAGGATCTAAGAGACATTAAGTTCTTGTGACTAAACCA 3134
 Db 2107 TTAGTAGAGGATCTAAGAGACATTAAGTTCTTGTGACTAAACCA 2154

RESULT 5
 BT004380
 LOCUS 1909 bp mRNA linear PLN 14-FEB-2003
 DEFINITION Arabidopsis thaliana clone U20243 unknown protein (At5g23570) mRNA,
 complete cds.
 ACCESSION BT004380
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 (bases 1 to 1909)
 Yamaoka, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,
 Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C.,
 Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,
 Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P.,
 Southwick, A., Tzipi, M.G., Wu, T., Shinzaki, K., Davis, R.W.,
 Ecker, J.R. and Theologis, A.
 Arabidopsis Open Reading Frame (ORF) Clones
 2 (bases 1 to 1909)
 Yamaoka, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,
 Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,
 Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,

TITLE
JOURNAL
COMMENT

Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C. J., Narusaka, M.,
Nguyen, M., Palm, C. J., Sakurai, T., Saitou, M., Seki, M., Shim, P.,
Southwick, A., Tripp, M. G., Wu, T., Shinzaki, K., Davis, R. W.,
Ecker, J. R. and Theologis, A.
Submitted (14-FEB-2003) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RFL CDNA's (RFL CDNA : 'RIKEN
Arabidopsis Full-Length CDNA') : Seki, M., Narusaka, M., Ishida, J.,
Saitou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Theologis, A.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RFL CDNA's: Yamada, K.,
Chan, M. M., Chang, C. H., Dale, J. M., Hsuan, V. W., Lee, J. M.,
Onodera, C. S., Quach, H. L., Tang, C. C., Toriumi, M., Wong, C., Wu, H. C.,
Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C. J., Nguyen, M.,
Palm, C. J., Shin, P., Southwick, A., Tripp, M. G., Wu, T., Davis, R. W.,
Ecker, J. R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP
/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome
submitted to Genbank.

FEATURES

source
Location/Qualifiers
1. 1909
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U20243"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1. 1909
/gene="At5g23570"
1. 1878
/gene="At5g23570"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="RA042374.1"
/db_xref="GI:28393933"
/translation="MSSRAGPMSKEKNVGSPTPEYVQVYAGTRILASSDQDGEW
EVSISKNNKRGNTSGKTWVSQNSNPRAWGQOQGRGSVNSRGNNVSGRGNGRG
IQANISGRGRALSRKYDNMFVAPPPVSRPRLBGMWQAKGSAQHTAVQEPFDD
VDNASEBENDSDALDSDDDLDASDDVDSQKSGSRKQNGKFKFPGSLDLSIEQ
INERPOQHPCACONGGATIDWYLIHPLHARTKARBYKRLRELAELVEKLDLQNG
ASTIPCEIYIGKQKGLGEDDEKDYEIWMPVITIMNRLDDNDKMLGNGNOELLFV
DKYELARHSYGRQGRGWSVLMFBSATGYIEAERHRELAEMGLDIRAMQKSM
FSGGVRLYGFPLATKQDLDFNCHSQKTKLELKSYSQMVKEVLRQISEDQQLAY
FKKLSKQNHAKVLESLIEIMSEKLRTEIDNRIVARPTKMHQENREMDHDF
MDSIKOIHERRDAKENFEMLOOQERAKVGGQOQINPSSNDCKRAEVSPTIEF
OEKEMEFVEREMILINDQEKMKEDMKRHRHEIFPLEKPEFDEALBQLMKHGLNED
D"

gene
1879..1909
/gene="At5g23570"

ORIGIN

3 UTR

Query Match 45.6%; Score 1493; DB 8; Length 1909;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 0; Indels 376; Gaps 4;
QY ATGAGTTCTAGGCTGTGCTCAATGCTTAAGAAAAAGAACGTTCAAGGTTGTTAAGGCT 755
DB 1 ATGAGTTCTAGGCTGTGCTCAATGCTTAAGAAAAAGAACGTTCAAGGTTGTTAAGGCT 60
QY 756 GAGGTTGAACGTTGTTCAAGGTTTGGCAGAGGACGAGACTGGCTTCTTCAACAAGAT 815
DB 61 GAGGTTGAACGTTGTTCAAGGTTTGGCAGAGGACGAGACTGGCTTCTTCAACAAGAT 120

QY 816 GGAGAGAGTGGAGGTCATTTCCAGAAGAACAAACCAAGAAAACCTTCTGGA 875
DB 121 GGAGAGAGTGGAGGTCATTTCCAGAAGAACAAACCAAGAAAACCTTCTGGA 180
QY 876 AAAAATTGGGTTTCTCAAGATTCGAATTCCTCTAGACTTGGGGTGTCTGACCAAGG 935
DB 181 AAAAATTGGGTTTCTCAAGATTCGAATTCCTCTAGACTTGGGGTGTCTGACCAAGG 240
QY 996 AGAGGTGACACGATATCTGGGAGAGAAACAATGTATCCGGAGAGGTAAACGCAATGT 995
DB 241 AGAGGTGACACGATATCTGGGAGAGAAACAATGTATCCGGAGAGGTAAACGCAATGT 300
QY 996 CGGGGCAATTCAGCTAACATATCTGTGGGAGACGAGCTTGGACAGAAATGATATAC 1055
DB 301 CGGGGCAATTCAGCTAACATATCTGTGGGAGACGAGCTTGGACAGAAATGATATAC 360
QY 1056 AACTTTGTGGCAACCCCACTGTATCTGCCCCCTCTTTGGAAGAGATGGAATTGGAG 1115
DB 361 AACTTTGTGGCAACCCCACTGTATCTGCCCCCTCTTTGGAAGAGATGGAATTGGAG 420
QY 1116 GCAAGAGAGGTTCTGCTCAGACACAGCTGTGACGAGATTCTCTGACGTGAGATAT 1175
DB 421 GCAAGAGAGGTTCTGCTCAGACACAGCTGTGACGAGATTCTCTGACGTGAGATAT 480
QY 1176 GTGATTAATGCTTCTGAGAAAGAAATGATTCGATGCTTGTGATGATTTGATGACAC 1235
DB 481 GTGATTAATGCTTCTGAGAAAGAAATGATTCGATGCTTGTGATGATTTGATGACAC 540
QY 1236 CTTCGAGATGATGATTAATGATCTCGGATGTGATGATTAATGATTAATGATTAAT 1295
DB 541 CTTCGAGATGATGATTAATGATCTCGGATGTGATGATTAATGATTAATGATTAAT 600
QY 1296 AATAAGTGGTCAAAAAGTCTTTGGCAGCTTGTGATGATGATGATGATGATGAT 1355
DB 601 AATAAGTGGTCAAAAAGTCTTTGGCAGCTTGTGATGATGATGATGATGATGAT 660
QY 1356 GAACCAAGAGGACAGTGCATGATTCGACCTGTGCAACCGAAGCTGTGATGATGAT 1415
DB 661 GAACCAAGAGGACAGTGCATGATTCGACCTGTGCAACCGAAGCTGTGATGATGAT 720
QY 1416 TTTAAGCTGACCCCTCTGAGCTCATGAGGACAAAGAGAGCTAGGAGATTAAGCTC 1475
DB 721 TTTAAGCTGACCCCTCTGAGCTCATGAGGACAAAGAGAGCTAGGAGATTAAGCTC 780
QY 1476 CATAGAGATTTGGCTGAATTTTAAAGAAAGATCTACAGATGAGAGGCGCATCTGCA 1535
DB 781 CATAGAGATTTGGCTGAATTTTAAAGAAAGATCTACAGATGAGAGGCGCATCTGCA 840
QY 1536 CCTTGTGTGAGATTTATGGGACGTGAAAGGCTTGGGTGAGATGAAAGATTAATGA 1595
DB 841 CCTTGTGTGAGATTTATGGGACGTGAAAGGCTTGGGTGAGATGAAAGATTAATGA 900
QY 1596 ATTGCTGGCCCTCAATGCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 1655
DB 901 ATTGCTGGCCCTCAATGCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 960
QY 1656 AAGGTGAATCTTCTGCTTACTTCTTAATTTCTTGAATTTCTTGAATTTCTTGA 1715
DB 961 A----- 961
QY 1716 GAATGTAATGTAATGAGCTGTGAGGCAACCAAGAGCTGTGAGATTTCTGACAA 1775
DB 962 -----AGTGGCTGGGCAATGAGGCAACCAAGAGCTGTGAGATTTCTGACAA 1007
QY 1776 GTATGAGCTCTTAAGGACAGGCAATTCATATGCTCCACAGGSCATCTGAGATGAGT 1835
DB 1008 GTATGAGCTCTTAAGGACAGGCAATTCATATGCTCCACAGGSCATCTGAGATGAGT 1067
QY 1836 TCTGATGTTGAGAGAGGTCACCTGGCTATTTGAGAGCCGAACGCTTCACCGGAGTT 1895
DB 1068 TCTGATGTTGAGAGAGGTCACCTGGCTATTTGAGAGCCGAACGCTTCACCGGAGTT 1127
QY 1896 AGCTGAGATGGGTTAGATGAATTCCTGGGCTCAGAAAGCCGACATATTTTCTGGAAG 1955

```

Db 1128 AGCTGAGATGGGGTTAGATAGAAATTCCTGGGGTCAGAAAGCGAGATGTTTCTCGAGG 1187
Qy 1956 TGTGGCCCACTGTATAGCTCTCTTCCAGCAAGAGCAATCTGACATATTCATCAACA 2015
Db 1188 TGTGGCCCACTGTATAGCTCTCTTCCAGCAAGAGCAATCTGACATATTCATCAACA 1247
Qy 2016 CTCTCAAGGTCTCTCCCCCAAGAAATTTGATATATGCTTTTATGTTTGTCAATGGAAT 2075
Db 1248 CTCTCA----- 1253
Qy 2076 TTAAGTTTGTGTGTCGGTTAATGATCTGTTATATATCTATGATTCATTAGG 2135
Db 1254 -----ACG 1256
Qy 2136 CAAAACAAAGCTGAATTTGAGTTGAATCATPACCAAGAGATGTTGTAAGAGCTGAG 2195
Db 1257 CAAAACAAAGCTGAATTTGAGTTGAATCATPACCAAGAGATGTTGTAAGAGCTGAG 1316
Qy 2196 GCAGATCTCTGAGCAATCAGACAGCTGAATCTTAAAGAACAGCTCTCAAAACAGAA 2255
Db 1317 GCAGATCTCTGAGCAATCAGACAGCTGAATCTTAAAGAACAGCTCTCAAAACAGAA 1376
Qy 2256 CAAGCAGCCAAAGTCTTGAAGAACTCTGAAATTTATGAGGAGAGCTGCTGAAC 2315
Db 1377 CAAGCAGCCAAAGTCTTGAAGAACTCTGAAATTTATGAGGAGAGCTGCTGAAC 1436
Qy 2316 TGCAGAGGATTAATCGATCTGTAGACAGAGAACTAAGATGACAGTAAACAGAGGGA 2375
Db 1437 TGCAGAGGATTAATCGATCTGTAGACAGAGAACTAAGATGACAGTAAACAGAGGGA 1496
Qy 2376 AGAGGATGATTTTCTAGAAATCAACAATTCGATTTGTATTAACCTACTGATTA 2435
Db 1497 AG----- 1498
Qy 2436 CATTTTGAATATATTCCTCAACAAAAACCTGTGTGTTGAAGATGATGACACGA 2495
Db 1499 -----AATGATGTCACACGA 1514
Qy 2496 CAGGTTTTCATGATTCATCAACAGATCCATGAAAGAGAGCGCAAGAGAGAGAA 2555
Db 1515 CAGGTTTTCATGATTCATCAACAGATCCATGAAAGAGAGCGCAAGAGAGAGAA 1574
Qy 2556 TTTTCGATGTTGACAGACAGAGAGAGTGTGGCAGACAGACAGAGAAAT 2615
Db 1575 TTTTCGATGTTGACAGACAGAGAGAGTGTGGCAGACAGACAGAGAAAT 1634
Qy 2616 TATCCCTCTAGCAATGACGATTCGGAAGAGGTATATGTACTAACAATATCCCT 2675
Db 1635 TATCCCTCTAGCAATGACGATTCGGAAG----- 1665
Qy 2676 CTGGGCTTTTGTTCCTAAACCTAAGAGTAATGATTCGGGTTTGAATCTTTGG 2735
Db 1666 ----- 1665
Qy 2736 CAGAGCTGAGAGATGTCAGAGCTTCATCGAGTTTCAAGAGAAAGAGATGAGAGTTGT 2795
Db 1666 -AGAGCTGAGAGATGTCAGAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGTTGT 1724
Qy 2796 GGAAGAGAGAGATGTCATTAAGATCAAGAGAAAGATGAGAGACATGAGAGAG 2855
Db 1725 GGAAGAGAGAGATGTCATTAAGATCAAGAGAAAGATGAGAGACATGAGAGAG 1784
Qy 2856 GCATCCAGAGAGATTTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAACAGTCA 2915
Db 1785 GCATCCAGAGAGATTTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAACAGTCA 1844
Qy 2916 GTACAAAGATGAGCTTCACAATGAGATGATGAGCAAAAGTCTGTACACAAGCAAG 2975
Db 1845 GTACAAAGATGAGCTTCACAATGAGATGATGAGCAAAAGTCTGTACACAAGCAAG 1904
Qy 2976 ACTAA 2980
|||

```

```

Db 1905 ACTAA 1909

RESULT 6
LOCUS AX078761 1878 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 2 from Patent WO0105951.
ACCESSION AX078761
VERSION AX078761.1 GI:13158380
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE
1 Becilín, C., Elmayan, T. and Vaucheret, H.
Novel sgss plant gene and use thereof
Patent: WO 0105951-A 2 25-JAN-2001;
AVENTIS CROPS SCIENCE S.A. (FR) ; INSTITUT NATIONAL DE LA RECHERCHE
AGRONOMIQUE (FR)

FEATURES
source
1..1878
location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
1..1878
CDS
1..1878
/translation="MSSRAGPWSKEKNGVGGVREPOLVQSLAGTRLLASDDGGEW
EYISKKNKRENTSGKTVVSNKSPRPAAAGQOQGRSSNVSGRNNVSGRNGRG
IOANTSGRRLSRKYDNNFVAPPPVSPLEGGNNWQARGSAAHTAVQEPVEDD
VDNAEENDSDALDDSDDLASDDYDSDVSGSKSRKQNKWFKFGSLDLSLEQ
INEPQRMHCPACONGPGLIDWYNLPLAHLARTGABRVKILHRELALELDLQMG
ASVPEGRITGQWGLGDEKDEKYEIVMPMTVIMNTRLDKDNKMLGNGQELQYF
DVEALRPARHSYGRQGHGSMVMESSATGYLELRHLRELAENGDLRIAMGQBSM
PSSGTRQLYGLATKODLDIRNHSQGRTRLEKLSIQEWVYKELROISENQOLNY
FNKLSKQNKAKVALESELEIMSEKLRATLAEENRIVRQTKQKHONREMDAHRFE
MDSIQIHERRPAKNEFEMLQQQRAVVGQOQONINPSSNDCKRAEVSFIEF
QEKEMEFVEEREMLIKQEKEMEDMKRHHBEIPDLKEPDEALEQMLKGLHNE
D"

ORIGIN
Query Match 44.6%; Score 1462; DB 6; Length 1878;
Best Local Similarity 83.3%; Pred. No. 4e-311;
Matches 1878; Conservative 0; Mismatches 0; Indels 376; Gaps 4;

Qy 696 ATGAGTTCTAAGGCTGCTCCCAATGCTTAAGAGAAAGAGCTTCAAGGCTTATAGGCT 755
Db 1 ATGAGTTCTAAGGCTGCTCCCAATGCTTAAGAGAAAGAGCTTCAAGGCTTATAGGCT 60
Qy 756 GAGTTGAACAGTTGTTCAAGTTTGGCAGAGGAGCAGACTGGCTTCTTCAACAATGAT 815
Db 61 GAGTTGAACAGTTGTTCAAGTTTGGCAGAGGAGCAGACTGGCTTCTTCAACAATGAT 120
Qy 816 GGAGAGAGTGGAGGCTATTTCCAGAGAAACAAGAACAAACCAAGAAACACTTCTGGA 875
Db 121 GGAGAGAGTGGAGGCTATTTCCAGAGAAACAAGAACAAACCAAGAAACACTTCTGGA 180
Qy 876 AAAACTGGGTTCTCAGAAATTCGAATCCTCTAAGAGCTGGGGTGTGACAGAGAGGG 935
Db 181 AAAACTGGGTTCTCAGAAATTCGAATCCTCTAAGAGCTGGGGTGTGACAGAGAGGG 240
Qy 936 AGAGGTAGCAAGTATCTGAGAGAGAAACAATGTATCCGAGAGAGTAAACGCGCAATGT 995
Db 241 AGAGGTAGCAAGTATCTGAGAGAGAGAAACAATGTATCCGAGAGAGTAAACGCGCAATGT 300
Qy 996 CGGGGCATTTCAAGCTPAACATATCTGCTGGGGAGCAGAGCGTTGAGCAGAAAGTATGATAAC 1055
Db 301 CGGGGCATTTCAAGCTPAACATATCTGCTGGGGAGCAGAGCGTTGAGCAGAAAGTATGATAAC 360

```


TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	COMMENT
lepiniec, U., Caboche, M. and Lecharny, A.					
T-DNA integration into the Arabidopsis genome depends on sequences					
of pre-insertion sites					
EMBO Rep. 3 (12), 1152-1157 (2002)					
22363535					
12446565					
2 (bases 1 to 650)					
Balzerque, S.					
Direct Submission					
Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue					
Gaston Cremieux, 91057 Evry cedex, FRANCE					
PCR was performed on DNA from transformants of Arabidopsis thaliana					
plants from INRA (Versailles). The DNA fragment(s) resulting from					
the PCR were directly sequenced from the left or the right border					
to determine the genomic sequence flanking the insertion. T-DNA					
derived sequences were removed. Information to order the					
corresponding mutant line and a link to a database providing a					
graphical display of the insertion site are available at					
http://dbsgap.versailles.inra.fr/publicines/. This sequence has					
been generated in the framework of the French plant genomes					
program 'Genoplatane' (http://www.genoplatane.com and					
http://genoplatane-info.infobiolegen.fr).					
Location/Qualifiers					
1..650					
/organism="Arabidopsis thaliana"					
/mol_type="genomic DNA"					
/cultivar="Wassiljewskij1a"					
/db_xref="taxon:3702"					
/clone="157C08"					
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"					
1..650					
/note="T-DNA flanking sequence					
left border"					
misc_feature					
19.7%; Score 643.6; DB 8; Length 650;					
Best Local Similarity 99.4%; Pred. No. 6.9e-131;					
Matches 646; Conservative 0; Mismatches 4; Indels 0; Gaps 0;					
Query Match					
1428 CCTCTACTAGCTCATGCGAGACAAAGAGACTAGCGAGTTAAGCTCCATAGAAATTG					
1 CCTCTACTAGCTCATGCGAGACAAAGAGACTAGCGAGTTAAGCTCCATAGAAATTG					
1488 GCTGAGGTTTAAAAAGATCTACAGATGAGAGGCGCATCTGTCATTCTTGTTGAG					
61 GCTGAGGTTTAAAAAGATCTACAGATGAGAGGCGCATCTGTCATTCTTGTTGAG					
1548 ATTATATGGGCGTGGAAAGGTTTGGGTGAGAGATGAAAAGATTATGAAATTGTCGGCT					
121 ATTATATGGGCGTGGAAAGGTTTGGGTGAGAGATGAAAAGATTATGAAATTGTCGGCT					
1608 CCAATGGTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG					
181 CCAATGGTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG					
1668 TTCTGTCTTACTCTTTAAATTTTCTCTTGATCTTACTGATCTTAAATGTTACTT					
241 TTCTGTCTTACTCTTTAAATTTTCTCTTGATCTTACTGATCTTAAATGTTACTT					
1728 GTATGGCTCGGCAATGGGCAACCAAGAGCTGCTGAATATCTTCCACAATGATGAGCTCT					
301 GTATGGCTCGGCAATGGGCAACCAAGAGCTGCTGAATATCTTCCACAATGATGAGCTCT					
1788 TAGAGCAAGCAATTCCTATGTTGCAAGAGGCGCATGCTGATGATGATGATGATGATG					
361 TAGAGCAAGCAATTCCTATGTTGCAAGAGGCGCATGCTGATGATGATGATGATG					
1848 GAGCAGTGCCACTGCTGATTTTGAAGCCGAACGCTCCACCGGAGATTAGCTGATGGG					
421 GAGCAGTGCCACTGCTGATTTTGAAGCCGAACGCTCCACCGGAGATTAGCTGATGGG					
1908 GTTAGATGAAATTCCTGAGGCTCAGAAAGCAATGATTTTCTGAGGATGTTCCCACT					

Db		481	GTTAGATAGAATTCTCGGGGTCAGAAAGCCAGATGTATTTTTCTGAGGTGTGTCACCAACT	540
Oy		1968	GTATGGCTTCTTGCAACGAAGCAGACTCTGCACATATTCATTAACAAGCTTCAAGTTC	2027
Db		541	GTATGGCTTCTTGCAACGAAGCAGACTCTGCACATATTCATTAACAAGCTTCAAGTTC	600
Oy		2028	TCTCCCCAAGAATAATTGATATATGCTTTTAGTTTTGCTATTCGGAATTT	2017
Db		601	TCTCCCCAAGAATAATTGATATATGCTTTTAGTTTTGCTATTCGGAATTT	650
RESULT 8				
AKO64217				
LOCUS				
DEFINITION	AKO64217	1556 bp	mRNA	linear
				PLN 24-JUL-2003
ACCESSION				
VERSION	Oryza sativa (japonica cultivar-group)		cDNA clone:002-104-F07,	Full
KEYWORDS	insert sequence.			
SOURCE	AKO64217			
ORGANISM	FLJ_CDN1: oligo-capping.			
	Oryza sativa (japonica cultivar-group)			
	Oryza sativa (japonica cultivar-group)			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
	Pharbitidoideae; Oryzaceae; Oryza.			
REFERENCE				
AUTHORS	1			
	The Rice Full-length cDNA Consortium, National Institute of			
	Agrobiological Sciences Rice Full-Length cDNA Project Team;			
	Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,			
	Kishimoto,N., Yazaki,Y., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,			
	Kojima,K., Namiki,T., Ohneda,E., Yabagi,W., Suzuki,K., Li,C.,			
	Ohtsuki,K., Shishiki,T., Foundation of Advancement of International			
	Science Genome Sequencing & Analysis Group; Otsuo,Y., Murakami,K.,			
	Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,			
	Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,			
	Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Nikura,J.,			
	Iweda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,			
	Kusunegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;			
	Kawai,J., Carninci,P., Adachi,U., Aizawa,K., Arakawa,T., Fukuda,S.,			
	Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,			
	Kagawa,I., Kondo,S., Kono,H., Miyazaki,A., Otsuo,N., Oca,Y.,			
	Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,			
	Yoshino,M. and Hayashizaki,Y.			
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from			
	japonica rice			
JOURNAL	Science 301 (5631), 376-379 (2003)			
MEDLINE	22752273			
PUBMED	12869764			
REFERENCE	2 (bases 1 to 1556)			
AUTHORS	Adachi,U., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,			
	Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,			
	Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiroaka,T.,			
	Hori,F., Hotta,I., Iida,Y., Iida,Y., Iweda,R., Inamura,K.,			
	Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,			
	Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,			
	Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,			
	Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Kono,H., Kouda,M.,			
	Koya,S., Kuwahara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,			
	Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,			
	Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,			
	Namiki,T., Narikawa,R., Nikura,J., Nishi,K., Oka,M., Ooka,H.,			
	Numaeki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,			
	Otsuo,N., Oca,Y., Otsuo,Y., Ryu,R., Satoh,H., Sakai,C., Sakai,K.,			
	Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,			
	Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,			
	Shinagawa,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,			
	Tegawa,A., Takahashi,F., Takeuchi-Akita,S., Tanaka,T., Tomaru,A.,			
	Yoda,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yabagi,W.,			
	Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and			
	Yoshimura,A.			
TITLE	Direct Submission			
JOURNAL	Submitted (05-DEC-2001) Shoichi Kikuchi, National Institute of			
	Agrobiological Sciences, Department of Molecular Genetics, Head of			
	Laboratory of Gene Expression; 2-I-2 Kamondai, Tsukuba, Ibaraki			


```

exon      /number=13
           complement(24024..24102)
           /gene="T14E10_10"
intron    /number=14
           complement(24103..24708)
           /gene="T14E10_10"
exon      /number=14
           complement(24709..24855)
           /gene="T14E10_10"
intron    /number=15
           complement(24856..25105)
           /gene="T14E10_10"
exon      /number=15
           complement(25106..25345)
           /gene="T14E10_10"
intron    /number=16
           complement(25346..25617)
           /gene="T14E10_10"
exon      /number=16
           complement(25618..25747)
           /gene="T14E10_10"
gene       /number=17
           27262..30665
           /gene="T14E10_20"
           join(27262..27331,27416..27633,27731..27982,28309..28586,
           29667..30057,30207..30665)
           /gene="T14E10_20"
           /note="strong similarity to LeOPT1 oligopeptide
           transporter, Lycopersicon esculentum, EMBL:AF016713
           contains PTR2 family proton/oligopeptide symporters
           signatures AA73-97"
           /codon_start=1
           /product="oligopeptide transporter-like protein"
           /protein_id="CAB7565.1"
           /db_xref="GI:7258348"
           /db_xref="SPTREMBL:O9M112"
           /translation="MADSTSLINKRTKGGNNALFIIVEIAERFAFYGLASNLITFL
           TNEIGOSTATAKNITWTIGVSCMPFLIGAFIADSLIGPKYLLISFYLGITWLP
           LSVTVARRRREKVFPMALYVAVGEGHKPCMTTRADQGEANAEKAATSPFNY
           WMAIVLASSIAVIALIFIOERWSLIGSIIAGSVIAIVIFLIGIPYRKQVPPVGS
           PFTRAVQVVAALKKWRLSTRHYGLCYEEDEHLESTNSQVYLARTNOFREL"

Query Match      5.6%; Score 181.8; DB 8; Length 94911;
Best Local Similarity 74.8%; Pred. No.2.8e-29;
Matches 291; Conservative 0; Mismatches 72; Indels 26; Gaps 4;

QY 2045 TGATATATGCTTTAGTTTGTTCATTGGAATTTAAAGTTTGTGTCGCTGTTAATGCA 2104
    |||||
DB 57198 TCATTAACATTTCTAGTCCTTCATGCGATTTAGAGTTTGTGTTCTGTGAAATGCA 57139
    |||||
QY 2105 TCTGTTATGATATATCTATGATTCATTAGGCAAAACAGGCTGAAATTCGATTTGAAT 2164
    |||||
DB 57138 TCA--TATGATATTTAGTATGATTTACTCTATGCGCAAAAGGCGTGAATTTGAGATTAAT 57081
    |||||
QY 2165 CATACCAAGATGCTGTTAAAGAGCTGAGCAGATCTCTGAGCAATCAGACAGCTG 2224
    |||||
DB 57080 CTTACCAAGAAATGCTGTTAACTAGCTAGGCAATCTCTGAGCAATCAACACTGA 57021
    |||||
QY 2225 ACTACTTTAAGAACAAAGCTCTCAAAAAGAACAGACAGCCAAAGCTGCTTGAGAAATCTC 2284
    |||||
DB 57020 TCTGTTACAGAACAAAGATCACAAATACGAAG-----AGTATGGCTCTA 56974
    |||||
QY 2285 TGGAAATTTAGCGGAGAAAGCTCGTGAAGTCTGACAGATTAATCGGATCGTAGACAGA 2344
    |||||
DB 56973 TGGAAATTTAGTGTGAATCTGCGTGAACCTGCAAGACATA-----ATCGTAGACCA 56919
    |||||
QY 2345 GAACCTAAGATGCGAGTGAACAGACAGGGAAGAGGATGATTTTCC-----TAGAA 2398
    |||||
DB 56918 GAACCAAGATTCACATGACAGACAGAGAGAGGTAAGTTATTTCCGAAAAAATA 56859
    |||||
QY 2399 ATCACAACTGACATTTGTATTAACCTA 2427
    |||||

```

```

DB 56858 CATACCAACTGACATTTGTATTAACCTA 56830

RESULT 11
CNS08CAO/c
LOCUS
DEFINITION
Oryza sativa chromosome 12, . BAC OJ1561 A05 of library Monsanto
from chromosome 12 of cultivar Nipponbare of sep. japonica of Oryza
sativa (rice), complete sequence.

ACCESSION
AL831811.4 GI:28892661
VERSION
HTG.
KEYWORDS
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriatroidae; Oryzae; Oryza.
REFERENCE
1 (bases 1 to 91053)
Choisne,N., Orjeda,G., Catolico,L., Demange,N., Wincker,P.,
Segurens,B., Pelletier,E., Scarpelli,C., Salanoubat,M.,
Weissenbach,J. and Quetier,F.
Oryza sativa chromosome 12 sequencing
Unpublished
2 (bases 1 to 91053)
Genoscope.
Direct Submission
Submitted (20-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Mar 9, 2003 this sequence version replaced gi:28460590.
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: sequef@genoscope.cns.fr

-----
The following sequence is oriented from the T7 to the SP6 end. The
nucleotide sequence of this BAC clone was generated by combining
Monsanto, Syngenta and Genoscope sequencing data.
Upstream BAC (overlapping the T7 end) : OSJNBa0029N15 (AC-BX536967)
Downstream BAC (overlapping the SP6 end) : OJ1102_B11 (AC=AL713901)
FINISHED SEGMENT STARTS AT BASE 1
FINISHED SEGMENT ENDS AT BASE 75085
-----

FEATURES
source Location/Qualifiers
1..91053
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="12"
/clone="OJ1561 A05"
/clone_1fb="Monsanto"

ORIGIN

Query Match      5.3%; Score 174; DB 8; Length 91053;
Best Local Similarity 60.8%; Pred. No.1.4e-27;
Matches 320; Conservative 0; Mismatches 200; Indels 6; Gaps 2;

QY 1155 TTTCCTGACGTGAGAGATGATGATATGCTTCTGAGGAGAGAAATGATCCAGTCT 1214
    |||||
DB 81521 TTGTCTGTTGATGATCTGAAACAGACAAATGCTGAAGTTAAGACACATCGGACGATGT 81462
    |||||
QY 1215 TTGATGATTTGATGACGACCTTGCAAGTATGATATGATCTGAGATGATGCAAAAG 1274
    |||||
DB 81461 AATGATGATGATATGATGACATTTGAGTATGATGATTTGATGATGACATGAGAAA 81402
    |||||
QY 1275 AGCCATGATCAGGAAGAGATAAGTGTCAAAAAGCTTTTGGCAGCTTTGATATAG 1334
    |||||
DB 81401 AGTTTGGAGACTCGGAAAAAATACAAAGTTTCCAAAGTTTCTTGAAGTCTCTGAGGCG 81342
    |||||
QY 1335 TTGTGATGACAGACATTAATGAACACAGAGGAGTGGCATTTGCCAGCTTGAGAAC 1394
    |||||

```


TITLE

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

JOURNAL
MEDLINE
PUBMED
AUTHORS

2 (bases 1 to 2570)

12869764
22752273

Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imanura, K., Imocani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toy, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, M., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE
JOURNAL

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28k full-length cDNA clones from japonica rice.

COMMENT

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imanura, K., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toy, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

Location/Qualifiers

1..2570

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="U023114M02"

ORIGIN

Query Match 4.8%; Score 155.8; DB 8; Length 2570;
Best Local Similarity 54.3%; Pred. No. 1,76-23;
Matches 460; Conservative 0; Mismatches 307; Indels 80; Gaps 4;

Qy	1172	TGATGTGATTAATGCTTCTTGAGAGAGAGATATTCCTGCTTTGGATGATTCGATGA	1231
Db	713	TGAAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	772
Qy	1232	CGACCTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1291
Db	773	TGATGATACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	832
Qy	1292	GCAGATTAAGTGTTCATTAAGTGTTCATTAAGTGTTCATTAAGTGTTCATTAAGTGTTCAT	1351
Db	833	AAGGACCAAGCTGTTGAAGATTTTGAATCTTCGAGAGAGTTAAGTGTTCATTAAGTGTTCAT	892
Qy	1352	AAATGAACCAAG	1411
Db	893	AAATGAACCAAG	951
Qy	1412	TTGATTAATA---CTGCACCTCTACTAGCTGATGATGATGATGATGATGATGATGATGATGAT	1468
Db	952	CTGATACAAAGGGTTAAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1011
Qy	1469	TGAGTTCATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1528
Db	1012	TGAGTTCATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1071
Qy	1529	TGTCATTCCTTGTGTGATTAATGAG	1588
Db	1072	AATGCAATATCTGCGAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1128
Qy	1589	TGATGAATTCCTGCGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1648
Db	1129	TGCTCTTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1188
Qy	1649	TGACATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1708
Db	1189	GGATGGA-----AGTGAAAGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1196
Qy	1709	GATCTTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1768
Db	1197	-----AGTGAAAGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1235
Qy	1769	TGCAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1828
Db	1236	TGATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1295
Qy	1829	TGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1888
Db	1296	TGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1355
Qy	1889	GGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1948
Db	1356	ATCACCTTGTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT	1415
Qy	1949	CTGAGAGTGTTCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2008
Db	1416	CTGAGAGTGTTCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1475
Qy	2009	ATCAACA 2015	
Db	1476	ATGACA 1482	

RESULT 14

AF542974

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

AF542974 1564 bp mRNA linear PLN 02-JUL-2003

Triticum aestivum Emri mRNA, complete cds.

AF542974.1 GI:32401385

AF542974.1 GI:32401385

Triticum aestivum (bread wheat)

ORGANISM *Triticum aestivum*
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
AUTHORS 1 (bases 1 to 1564)
TITLE Zhao, X., Li, Q. and Zhang, X.
JOURNAL Isolation and expression of a new kind of gene involved in
AUTHORS embryogenesis in *Triticum aestivum* L.
TITLE Unpublished
2 (bases 1 to 1564)
REFERENCE Zhao, X., Li, Q. and Zhang, X.
AUTHORS Direct Submission
JOURNAL Submitted (03-SEP-2002) College of Life Sciences, Shandong
Agricultural University, Daizong Street, Tai'an, Shandong 271018,
P.R. China

FEATURES
source Location/Qualifiers
1..1564
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultiivar="PH 82-2-2"
/db_xref="taxon:4565"
/dev_stage="embryo"
309..1223
/note="involved in embryogenesis"
/codon_start=1
/product="Hmrl"
/protein_id="AAP80862.1"
/db_xref="GI:32401386"
/translation="MSGNOELLEYSDYATKARHAYPGHGRMSVLIPESSAVGME
AERLHKHPIIDORTDTONRVRVPLPGKRLQYGLARKEDEMPNPHCOGSRKY
EMRSHNEVVAQKMSQEDNQALYLNKRVKTEORSKVEETLGLVITOKLRETMEN
IIVRSKAEKSEEEEMKSEKFEFHDIINIHATKEDSKSEFERLQSRARACCD
VDSGTENRRLRKEOVORFICOVDOVEFAERDEMIKAEKVKQLKKEVMAKVE
LEKEDPALITGLMEKHRGTQASSSP"

CDS
ORIGIN

Query Match 3.7%; Score 120.8; DB 8; Length 1564;
Best Local Similarity 61.2%; Pred. No. 8.8e-16;
Matches 213; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

1717 AATGTTACATTTGATGCTGGCTGGGCAACCAAGAGCTGCTGATCTTGACAG 1776
285 ATGATGCTCAGTCAAGTGTGATGAGCATGAGCAACCAAGAACTTCTGATATTTCAAGTAT 344
1777 TATGAGGCTCTTGAAGCAGCCATTCCTATGCTCAAGGCCCATGCTGGATGAGTGT 1836
345 TATGCTGCAACCAAGCAGCCATGATGCTCAGGTGGCACCGTGGCATGAGTGTG 404
1837 CTGATGTTTGAAGCAGTGCATCTGCTATTTGGAGCGCAAGCGCTCCACCGGAGTTA 1896
405 CTAAATATTGAAGCGCTGCTGTGGGCTATATGAGGCAAGACGTTCTCATAGCATTTT 464
1897 GCTGAGATGGGGTTAGATAGTAATTTGCTG---GGGTCAAGAGCGCATGTTTCTGGA 1953
465 ATTATATCAAGAAAGACAGGACACTTGGCAGAAATCGCAGGGTTCCTTTCTTACCTGAT 524
1954 GGTGTGCGCAACTGTATGCTTCTTGGCAAGCAAGCAAGATTTGACATATTCATCA 2013
525 GGGAGAGACAAATATACGTTTCTTGAAGCAAGAAAGACATGAGACTTTCAATAGA 584
2014 CACTCTCAAGGTTCTCTCCCAAGAAATTTGATATATCTTTTACT 2061
585 CACTGCCAAGGAAAGCGCGCTGAAATACAGATGAGATCTCATAT 632

RESULT 15
AF469493 513 bp mRNA linear PLN 02-JUL-2003
LOCUS Triticum aestivum H0TR mRNA, partial cde.
DEFINITION AF469493
ACCESSION AF469493.1 GI:32400755
VERSION
KEYWORDS
SOURCE Triticum aestivum (bread wheat)

ORGANISM *Triticum aestivum*
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
AUTHORS 1 (bases 1 to 513)
TITLE Li, J.R., Wang, F., Li, Q.Z. and Zhang, X.S.
JOURNAL Gene isolation and expression of a new Zn-finger
AUTHORS Unpublished
2 (bases 1 to 513)
REFERENCE Li, J.R., Wang, F., Li, Q.Z. and Zhang, X.S.
AUTHORS Direct Submission
JOURNAL Submitted (18-JAN-2002) College of Life Science, Shandong
Agricultural University, Dai zong Street 61, Taitan, Shandong
271018, P.R. China

FEATURES
source Location/Qualifiers
1..513
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultiivar="PH 82-2-2"
/db_xref="taxon:4565"
/dev_stage="seeds"
1..513
/note="similar to H0TR of Hordeum vulgare"
/codon_start=1
/product="H0TR"
/protein_id="AAP80610.1"
/db_xref="GI:32400756"
/translation="WLGNOELLEYSDYATKARHAYPGHGRMSVLIPESSAVG
YMAERLHKHPIIDORTDTONRVRVPLPGKRLQYGLARKEDEMPNPHCOGSR
IKYEMRSHNEVVAQKMSQEDNQALYLNKRVKTEORSKVEETLGLVITOKLREIM
ENIFVRSKA"

CDS
ORIGIN

Query Match 3.6%; Score 117; DB 8; Length 513;
Best Local Similarity 61.6%; Pred. No. 6.3e-15;
Matches 205; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

1732 TGGCTGGCATGGGCAACCAAGAGCTGCTGAAATCTTGACAAAGATGAGGCTTTAGA 1791
1 TGGTTAGGCAATGGGCAACCAAGAACTTCTTGATATTTCAAGTATGCTGGACCAAA 60
1792 GCACGCCATTCCTATGCTCAAGGGCCATGCTGGATGAGTGTCTTGATGTTTGAAGC 1851
61 GCACGCCATGATATGCTCCAGTGGCACCGTGCAAGTGTGCTAAATATTTGAAGC 120
1852 AGTGCATGCTGATTTGAGGCGCAAGCGCTCCACCGGAGTTAGCTGAGATGGGTTA 1911
121 TGTGCTGTGGCTATATGAGGCAAGCGTCTTCAATAGCATTTTATGATCAAGAACA 180
1912 GATGAATTCCTG---GGGTCAAGAGCGCATGATGTTTCTGAGGTGTTCCGCAACTG 1968
181 GACAGGACACTTGGCAGAAATCGCAGGGTTCCTTTCTTACCTGTGGAGAGACAAATTA 240
1969 TATGCTTCTTGAAGCAAGCAAGATCTGCAATATTCATCAACTCTCAAGTTCT 2028
241 TAGGTTTCTTACCGCAAGAAAGACATGAGACTTTCAATAGACATGCGCAAGGAAA 300
2029 CTCGCCAAGAAATTTGATATATGCTTTTACT 2061
301 AGCGCTGAAATACAGATGAGATCTCATAT 333

Search completed: April 8, 2004, 13:33:35
Job time : 8568.15 secs

/codon_start=1
/evidence=not experimental
/protein_id="BAA97250.1"
/db_xref="GI:8809709"
/translation="MQPGTSVYSPALNQADEPPIISKTVVNDTKIIFPDGDADPL
PALENIQISGEYPGEHLOACGSIINGTSCNEFWCHLEDGSVNTIDAKKENYIVT
ADVQGLCLAEVQPLDDNRKQELVVFANDNRKLAICHEPMQSNIDIKTLHAGSKV
SLAIGFVHIMEAATLSIEREGYTIKCNNDLTTEKFSASTAVIKPEKPAELVITISD
GSEHCLRVNEMEDIRLCKEKKRLVLYVVVFEOKL"
complement(join(8642..8705,8794..8863,8978..9387,
9584..10347))
/note="gene_id:MOM1.22
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97249.1"
/db_xref="GI:8809708"
/translation="MSMRKSGSGWTAFLDKORQGLSEVEGDPEPVSIVNASF
GVGRGLRRNHESEKSSVLLPSPRPALTEKDCNQERGGCCRRKEDTLSPVNS
HDLAFTKLKEMNSWADNLIIDVLTSTEDFEALAFIKGVNSGKEDEPTKIEGY
SSDNRESEYRTPEKTVTSVKAABSTPEADGYLDENGSSFLVANAQNDKRPDI
SELDITIQLOSTPIPEWEEDDLVISHKDKALKRSASNSRAQNAFORDHASA
KCHSDVAREDMLAERKLAEEAAKIIIGITKONDIWKLDLHGLATEAAVQALQERQM
IGHFLPVNSVSPNCRSNAALRSASQEPFGLDEGWHQORTSRELRNLSIQVITG
ICKHSRQASLPLAVKTFPEDNRYRDETRPGVITVRPKFRHS"
complement(12417..13424)
/note="contains similarity to unknown protein
gb|AA27018.1
gene_id:MOM1.21"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97248.1"
/db_xref="GI:8809707"
/translation="NATDSQPNOKLTPLIKTRALTIVISTMDNQRPDCTINRRRL
LDFRAPNPKEVNIIVSTISDFVVDSDRIDMFRLYTHVSGDKI PVVFFHGGGFAFLS
PAAVYDNCRRPARKLPAYVISVNRVLAPEHRYPQYDGGDALQYIEENGLIIPA
NADLSRCPFAGSAGCNIHANVAIRICRPSRSTVAKIIGLISIQPFGEERTAE
KOLVGAIVSPDRDWCMAKGLNDRBEAVNAGPVNAVIDISGLDYETWVAVAGPDL
KOMQSYVEMLKCGKATLIEVPMNFHFIPELPEAGQLIMRIKDPVDERVASLS
A"
join(14922..14981,15072..15167,15563..15666,15755..15891,
16096..16188)
/note="gene_id:MOM1.20"
/codon_start=1
/evidence=not experimental
/product="50S ribosomal protein L24"
/protein_id="BAA97247.1"
/db_xref="GI:8809706"
/translation="MGMAAEKILRMKILRGDNWMIIRKDKGETGITIKRVISQR
VIVEGNLIKIKH KGGPDHGGIFVYEAQPLHASNVQVDPVTRGPKCVKYLEDDTK
VVAAGCTGSSGILRPDELIKIRATPRPTAGKDPMEFVWEQYDADATGKGMPL"
join(16823..16952,17143..17320,17412..17493,17675..17851,
17943..18110,18199..18390)
/note="gene_id:MOM1.19"
/codon_start=1
/evidence=not experimental
/product="26S proteasome, non-ATPase regulatory subunit"
/protein_id="BAA97246.1"
/db_xref="GI:8809705"
/translation="MERLQIRIFAGGGLGHASPDSPPLDTSQGVYISLALIKMLKHC
RAGVMEVGLMGEFVDEYIVVDFVAMPQSGTSVEAEDVHETITMLMLKOTG
RBMVVGWHSRFGCMLSGVDINTQGSFALNDAVAVVVDPIOSVAGKVVIAFR
SINPOTIMGCRPOTSNLGHKNPSIALIHGNAKRVAVQEBDELPEKLAIVANGR
LHKKWTDLIRRFDTSKINQVQEMSLAKAKNRKAVQEBDELPEKLAIVANGR
QAKKLEHVSNSNSNI VQTGLMLDVVF"
complement(join(18965..19043,19240..19238,19419..19536,
19666..19739,19908..19961,20387..20330))
/note="contains similarity to unknown protein
emb|CAA22897.1
gene_id:MOM1.18"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97245.1"

/db_xref="GI:8809704"
/translation="MDKNQALIEKMKLVGEVEDEQQADEESSLSFEMEDLRNCAL
TKORFYGALICLSAGLCTTSLMLVFNVPVFGTFTGLNMGALSTAFILGPOROV
TMMLDPAIRYATVLAIIIALFCALVYRNKLTLLIALLLEFTGLIWSLSIIPAR
TMVSKIRNCPTEP"
join(22672..22634,23708..23999,24111..24355,24458..24624,
24715..24925)
/note="emb|CAB62356.1
gene_id:MOM1.17
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97244.1"
/db_xref="GI:8809703"
/translation="MSSRAGPMSKEKNVQGYRPEVQVQGLAGTRLASQDDGGEW
EYISKKNKPGNTSGKTWVSONSPDPAAGCOQGRNMGQAGNNVSGRNGNGRG
IOANTSGRALSRTYDNNFVAPPPVSRPPELEGNNMQARGSAAHTAVQEPDYEDD
VMAEERNDSDALDDSDLDSDYSDVQSHSGSRKQNKPFKFSGLDLSIEQ
INEPORHNCPCQNGPAGIDWYNLHPILAHARTGARVKKLHRELAEVLEQDLMRG
ASVIFGEIYQMGKLGDEKDYELVMPPMVIMNTRLDKDNDKMLGQNGQELLEYF
DKVELAPRHSYQPGHGRGMSVLMFESSATGYLIEARLHRELAEVLEQDLMRG
BSGRIQLYGLATQDDIDIPNOSHQGRTRLEKSYQVEMVYKELROI SEDNOOLN
FNKLSKQNKAKVLESLETSERKLRTAEDNRLVRQTKMQHONREMDAHDFF
MDSIQIHERDRDAKENFEMLOQDERAVVQGOQONINPSNDCKRAEEVSSFIEF
DEKEMEVEVEEREMLIKQEKMEDEMKRHHHEIFDLKEFDEALQJLMYKRLGHNE
D"
join(25614..25838,26061..26207,26324..26438,26516..26610,
26710..26813,27522..27765,27943..28036,28115..28305,
28385..28604,28688..28855,28944..29092)
/note="gb|AAC97420.1
gene_id:MOM1.16
strong similarity to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97243.1"
/db_xref="GI:8809702"
/translation="MAPPAGQYAAVAEAGADGAPPOQGRGFTISGIVRAVWFY
ASKPSPKQMPDPAQOLMNTLPHKESLDWMFYLSQEFNFGNDRALYMETN
IPYAVTBEISRTKSLTYPSRTLLNNSSLYVHIFPARGPDIPTDPEYDPLNSPR
THAVATYPRKQKNNKSSILSPKSDSESEPEVEYQDKSPKEVPEWISIKKN
VTINLVDFPTQADLVEFTGNTYTFYFNFEKLRFDGFIIVNTSGLPLNLEISF
ISMWMOLEFOQVDSFQMRQSYGSMLEDSDEBILKVFLEDFVYLLGITMFVMSLVSF
DELAFKNDIQEWNKKNKMEGLSAKSVNLNFIQFPIFYLLDNDNSMMILASSGVGC
IEFKIKGAMRIEVDKSGMIPRLRPHDESVNSNTEKEDJAIKFLSVLLLVYGL
SIVSLAYBRKHSVMSIISLTSVYMGFTMCQPLFNMYLKSVAHLPMOMYKFE
LNTITDDLPAPVYIKRPIILHRLSVFPDDVIFLYLYQRNRYVPDKRVNFERGGEDET
AEKRLITEKEEEDKXTN"
complement(join(29439..29576,29661..29885,29985..30268,
30341..30493,30579..30722,30956..31484))
/note="gene_id:MOM1.15

Query Match 100.0%; Score 695; DB 8; Length 81365;
Best Local Similarity 100.0%; Pred. No. 7.8e-131;
Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CDS

QY 1 GCAAAACAAACAAATTAAGCAAGTCATGTCGTAGCAATAATTAATAGTGGACAA 60
DB 21977 GCAAAACAAACAAATTAAGCAAGTCATGTCGTAGCAATAATTAATAGTGGACAA 22036
QY 61 TTAAGTTAAGCAAAAGCAAAAGCAAAAGTCAAAATTAAGCAAAAGTCAAAAGTCAAT 120
DB 22037 TTAAGTTAAGCAAAAGCAAAAGCAAAAGTCAAAATTAAGCAAAAGTCAAAAGTCAAT 22096
QY 121 GAAATTTGAGTCCGAATCGGAAACAGAGCCGTTTAAAGCTTAATTAAGCTTCC 180
DB 22097 GAAATTTGAGTCCGAATCGGAAACAGAGCCGTTTAAAGCTTAATTAAGCTTCC 22156
QY 181 ATTGATCTCTTCTGTCAGTTATTTCTCTCGGAGTCTGATCACTACTCTCAC 240
DB 22157 ATTGATCTCTTCTGTCAGTTATTTCTCTCGGAGTCTGATCACTACTCTCAC 22216
QY 241 TCTCGGCGCTTAAACTTACCTCTCGGTCGTTACTGTAAATTTCTGCTTAAAG 300

Db 22217 TCTCCGCGCTTAAACTTACGCTTCCGTCCTTACTGTAAGTTTTCGCTTAAG 22276
 QY 301 CCTCCGATCGCTCAACCGCATCATCTGTGCTGCATTTCTTTTCTTGGCTGAAA 360
 Db 22277 CCTCCGATCGCTCAACCGCATCATCTGTGCTGCATTTCTTTTCTTGGCTGAAA 22336
 QY 361 ATTCGCTTAAATGTTCTGCATTTGCAAGGTTTGTGCTAAGGCTTACTTTTCCCTA 420
 Db 22337 ATTCGCTTAAATGTTCTGCATTTGCAAGGTTTGTGCTAAGGCTTACTTTTCCCTA 22396
 QY 421 TTTTAAAGTTCTTAGGTAACGATACCTGCGCTTACTGTTTGTGTAATTTGTGCT 480
 Db 22397 TTTTAAAGTTCTTAGGTAACGATACCTGCGCTTACTGTTTGTGTAATTTGTGCT 22456
 QY 481 TTACACGCTTAAATGTTCTGCATTTGCAAGGTTTGTGCTAAGGCTTACTTTTCCCTA 540
 Db 22457 TTACACGCTTAAATGTTCTGCATTTGCAAGGTTTGTGCTAAGGCTTACTTTTCCCTA 22516
 QY 541 TTGTTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 Db 22517 TTGTTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22576
 QY 601 ATGCGTTTGGCTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 660
 Db 22577 ATGCGTTTGGCTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 22636
 QY 661 ACTGCGGTTTCTTGGTGGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 695
 Db 22637 ACTGCGGTTTCTTGGTGGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 22671

RESULT 3
 LOCUS 166494 7218 bp DNA linear PAT 28-DEC-1997
 DEFINITION Sequence 14 from patent US 5670367.
 ACCESSION 166494
 VERSION 166494.1 GI:2724471
 KEYWORDS
 ORGANISM Unknown.
 SOURCE Unclassified.
 REFERENCE 1 (bases 1 to 7218)
 AUTHORS Dörner, F., Schefflinger, F. and Falkner, F. Gunter.
 TITLE Recombinant fowlpox virus
 JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
 FEATURES
 source 1..7218
 /organism="Unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 12.3%; Score 85.6; DB 6; Length 7218;
 Best Local Similarity 6.4%; Pred. No. 1.7e-07;
 Matches 25; Conservative 232; Mismatches 131; Indels 0; Gaps 0;
 QY 174 CTTCCTCATTTGCTCTCTCTCGCATGTTATTTCTTCCTCCGCGAGTCTGACTACTA 233
 Db 1094 YY 1153
 QY 234 CTCTCACTCTCCGCGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 293
 Db 1154 YY 1213
 QY 294 CTTAAGCTCCGATCGCTCAACCGCATCATCTGTGCTGCATTTCTTTTCTTGGCTG 353
 Db 1214 YY 1273
 QY 354 TGAATAATGCGCCATGTTCTGATTTGCAAGGTTTGTGCTAAGGCTTAACTTAACTTAA 413
 Db 1274 YY 1333
 QY 414 CCTTAAATTTAAATGTTCTTAAAGTAACGATCGCTGCTTAACTGTTTGTGCTAATTTG 473
 Db 473 YY

Db 1334 YY 1393
 QY 474 TTGTCCTTACCGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 533
 Db 1394 YY 1453
 QY 534 TTGCTTTTCTTCAATTAATGCGATT 561
 Db 1454 CTTAATCACTGATGATGATGATGAT 1481

RESULT 4
 LOCUS BT002944 2162 bp mRNA linear PLN 15-JAN-2003
 DEFINITION Arabidopsis thaliana clone RAFL14-93-K05 (R20243) unknown protein
 (A5G23570) mRNA, complete cds.
 ACCESSION BT002944
 VERSION BT002944.1 GI:27754622
 KEYWORDS
 ORGANISM Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 2162)
 YAMADA, K., CHAN, M. M., CHANG, C. H., DALE, J. M., HSUAN, V. W., LEE, J. M.,
 YU, G., YUAN, S., CARNINCI, P., CHEN, H., CHEUK, R., HAYASHIZAKI, Y.,
 ISHIDA, J., JONES, T., KAMIYA, A., KAWAI, J., KIM, C. J., NARUSAKA, M.,
 NGUYEN, M., PALM, C. J., SAKURAI, T., SATOU, M., SEKI, M., SHIMU, P.,
 SOUTHWICK, A., TRIPP, M. G., WU, T., SHINOZAKI, K., DAVIS, R. W.,
 ECKER, J. R. and Theologis, A.
 Arabidopsis Full Length cDNA Clones
 Unpublished
 2 (bases 1 to 2162)
 YAMADA, K., CHAN, M. M., CHANG, C. H., DALE, J. M., HSUAN, V. W., LEE, J. M.,
 ONODERA, C. S., QUACH, H. L., TANG, C. C., TORIUMI, M., WONG, C., WU, H. C.,
 YU, G., YUAN, S., CARNINCI, P., CHEN, H., CHEUK, R., HAYASHIZAKI, Y.,
 ISHIDA, J., JONES, T., KAMIYA, A., KAWAI, J., KIM, C. J., NARUSAKA, M.,
 NGUYEN, M., PALM, C. J., SAKURAI, T., SATOU, M., SEKI, M., SHIMU, P.,
 SOUTHWICK, A., TRIPP, M. G., WU, T., SHINOZAKI, K., DAVIS, R. W.,
 ECKER, J. R. and Theologis, A.
 Direct Submission
 Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
 Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 COMMENT

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M. M.,
 Chang, C. H., Dale, J. M., Hsuan, V. W., Lee, J. M., Onodera, C. S.,
 Quach, H. L., Tang, C. C., Toriumi, M., Wong, C., Wu, H. C., Yu, G.,
 Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C. J., Nguyen, M.,
 Palm, C. J., Shimu, P., Southwick, A., Tripp, M. G., Wu, T., Davis, R. W.,
 Ecker, J. R. and Theologis, A.
 Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)
 contributed equally to this work as PIs.
 Annotation based on July 2002 version of the Arabidopsis genome
 submitted to Genbank.

FEATURES
 source
 1..2162
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAFL14-93-K05 (R20243)"
 /ecotype="Columbia"

```

gene
5'UTR
CDS
/note="This clone is in a modified pBluescript vector3
(lambda ps) as a BamHI/XhoI insert."
1..2162
/gene="At5g23570"
1..91
/gene="At5g23570"
92..1869

```

/gene="At5g23570"
 /codon_start=1
 /evidence=experimental
 /product="unknown protein"
 /protein_id="AA022757.1"
 /db_xref="gi:27754623"
 /translation="MSSRACGPMSEKKNVQGGYRPEVQLVQGLAGTRILASSQDDGSEME
 EVLSKNNKPKGNTSGKTWTWSQNSNPRAWGQOQGGNSVSGGNNVSGGNGNGKGRG
 IOANIISGRGRALSRKYDNNFVAPPPVSRPLEGGMGNOARQGSAGHTVAOFFPVEYED
 VDNASEENSDALDDSDDDLSDYDSDYSDGSKHSGKKNKKWPKFFGSLDSTIFGQ
 INPQROBMHCACNGSGCALIDWVNLHLPLAARFKKARARVYLHSLAEVLEKDIQMGQ
 ASVIPCSEIYGQMKGLGDEKDYELVWPENYINNTLRDKDDNDKMLGAGQELLETFF
 DKYELARHRYGPGQHGKMSVLMFESSATYGLAEERLHAEVLMGLDRLAAGQKRM
 FSGGROLDYGLFTRKODLDIFNQHSQGTATFLPKLSQEVNVELDQISDQNOQLVYV
 FKNTLSKONKKAATVLESLIEMSEKLRRTAVDNIIVORTQMOHREENDADVFSTFR
 MDSIKTILSHERRDAKEENPEMLQOQERKVVQOQOQNIIPSSNDICORAEVSSATFRF
 QEKMEBFVEREMLIKDQEKMEDMKKRHHIEFLDEKEDBALBQLMYKGLHNEED
 D"
 1970. 2162
 /gene="At5g23570"
 2146
 /gene="At5g23570"
 /note="not present in genomic sequence"

[illegible]

1651 bp contig from 38821 to 40471
66410 bp contig from 40572 to 106991
16357 bp contig from 107082 to 123438
2460 bp contig from 123539 to 125998
20897 bp contig from 126099 to 146595
37440 bp contig from 147096 to 184535

```

Overall quality chart :
Range      : bases
0          : 1377
1 - 9     : 1678
10 - 19   : 1569
20 - 29   : 3627
30 - 39   : 13450
40 - 49   : 13488
50 - 59   : 10047
60 - 69   : 22169
70 - 79   : 54644
80 - 89   : 43202
90 - 99   : 19284
-----
Percentage of bases with a quality value >= 40 : 88 %.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

```

[illegible]

[illegible]

LOCUS	AX344566	349980 bp	DNA	linear	PAT 01-FEB-2002
DEFINITION	Sequence 17 from Patent WO0200932.				
ACCESSION	AX344566				
VERSION	AX344566.1 GI:18492452				
WORDS	.				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1				
AUTHORS	Olek,A., Piepenbrock,C. and Berlin,K.				
TITLE	Diagnosis of known genetic parameters within the mhc				
JOURNAL	Patent: WO 0200932-A 17 03-JAN-2002;				
FEATURES	Epigenomics AG (DE)				
Source	Location/Qualifiers				
	1..349980				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="Chemically treated genomic DNA (Homo sapiens)-Original length of seq 1: 3.673778 <223>-split as follows:-seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001 649.980-seq 03 600.001 949.980-seq 04 900.001 1.249.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001 1.849.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001 2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001 3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001 3.649.980-seq 13 3.600.001 3.673.778 <223>-Original length of seq 2: 3.673778 <223>-split as follows:-seq 14 0.000.001 TO 0.349.980-seq 15 0.300.001 649.980-seq 16 600.001 949.980-seq 17 900.001 1.249.980-seq 18 1.200.001 1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001 2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001 2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001 3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001 3.673.778"				

ORIGIN

Query Match	11.7%;	Score 81.2;	DB 6;	Length 349980;
Best Local Similarity	46.0%;	Pred. No. 7.6e-07;		
Matches 275; Conservative	0;	Mismatches 323;	Indels 0;	Gaps 0;

QY	80	AAAAAAAAAGGTCAAAAATGAAACAAATCAATCGATGATAAAATTGGAGTCCAGAA	139
Db	71449	AAAAAAAAAAAAAAAAAAAAAAAAAATTAATGATGAAAGGTATATATTTTTAT	71508
QY	140	TCGAAAAACGAGCCGTTTAGACTTATATAAGCTTCTCATTTGTCCTTCTTCGTCA	199
Db	71509	AAAGTAATATGAAATAGTAATATACGCTTTTTTTTTTTTTTTTTTTTTTTTTTT	71568
QY	200	GTTATATTTCTCTCCGAGTCTCTACATCACTCTCACTCTCCGGGCTTAACTT	259
Db	71569	TT	71628
QY	260	ACGTCCTCCGCTTACTCTGTAAATTTTCTGCTTAGAGCCTCGATGCGCTCACCGC	319
Db	71629	TT	71688
QY	320	ATGCATCTGTCGTCAATTTCTCTTTTCTGCGTGAATAATGCCATATGTTCTCGA	379
Db	71689	TT	71748
QY	380	TTTCGAAGTTTTTGTCATAGGCTTACTTTTCCCTAATTTTATAGTCTTAGCTAA	439
Db	71749	TTTTTATTT	71808
QY	440	CGACACTGCGCTTACTGTTTTGTTGTCATTTGTTGTCGCTTACACGTTTAGTGCCTGA	499
Db	71809	TT	71868
QY	500	TCGAGATATTCATCTGAAAAAATCCTTCGTTTTTGTGTTTCAATATAAACGGA	559
Db	71869	TTTTTTTTTTTTTTTTTTTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	71928
QY	560	TTGATCTACCTTTTGTCCTTGATGTTGTGTTTGAAGCCATAGCGTGTGACTTGTTA	619
Db	71929	TTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	71988
QY	620	TAACTCAGCTTCAATGTCGGAATTTGAAATTTGGTATGACTGAGGTTCTTTGG	677
Db	71989	TTTTTTTGTGTTTTTTTTTTTTTTTTTTTTTTTGGAGATAGAGATTTGTTGTATAG	72046

RESULT	7
AJ592026	
LOCUS	
DEFINITION	AJ592026 1407 bp DNA linear PLN 23-OCT-2003
ACCESSION	
VERSION	AJ592026
KEYWORDS	right border; T-DNA flanking sequence.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1
REFERENCE	
AUTHORS	Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Craund, C., Deposse, R., Pelletier, G., Leplintec, L., Caboche, M. and Lecharny, A. T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites EMBO Rep. 3 (12), 1152-1157 (2002)
TITLE	
JOURNAL	
MEDLINE	
PUBMED	22363535
REFERENCE	12446565
AUTHORS	2 (bases 1 to 1407) Balzergue, S.
TITLE	Direct Submission
JOURNAL	Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue

COMMENT

Caston Creteux, 91057 Evry cedex, FRANCE
PCR was performed on DNA from transformants of *Arabidopsis thaliana* plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA⁺-derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.inbio.gen.fr>).

FEATURES

Source

```

1..1407
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Naasillewskja"
/db_xref="taxon:3702"
/clone="600D05"
/clone_1b="Arabidopsis thaliana T-DNA insertion lines"
1..1407
/note="T-DNA flanking sequence"
right border

```

ORIGIN

Query Match	10.5%;	Score 72.8;	DB 8;	Length 1407;
-------------	--------	-------------	-------	--------------

[illegible]

RESULT 8	1453 bp	DNA	linear	PLN 23-OCT-2003
AU591978				
LOCUS				
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence, right border, clone			
ACCESSION	558D01.			
VERSION	AU591978.1	GI:37941602		

KEYWORDS
SOURCE
ORGANI

right border: T-DNA flanking sequence.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; *Arabidopsis*.

REFERENCE

Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.

Lepiniec, L., Cabocne, M. and Lecharny, A.

॥॥॥॥

EMBO Rep. 3 (12), 1152-1157 (2002)

MEDLIN

12446565

REFERENCE

Balzerque, S.

TITLE
TOPICS

Gaston Cremieux 91057 Eyry cedex FRANCE
Submitted (23-Oct-2003) Balzeigues s.; UNKOV, INNA/CNRS, 2 rue

COMMENT

Plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsnp.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante.info.infobio.gen.fr>).

FEATURES

Source

```

1. .1453
  /organism="Arabidopsis thaliana"
  /mol_type="genomic DNA"
  /cultivar="Wassilewskaja"
  /db_xref="taxon:3702"
  /clone="598D01"
  /clone_1b="Arabidopsis thaliana T-DNA insertion lines
1. .1453
  /note="T-DNA flanking sequence
right border"

```

ORIGIN

Query Match	10.18;	Score 70;	DB 8;	Length 1453;
-------------	--------	-----------	-------	--------------

QY	157	TTTTAGAGCTAATAAGACTCCATATGTCCTCTCGACATTATTTCTCTCC	218
Db	331	TTTTTNTTTTNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	390
QY	217	GGAGTCCGACTCACACTCTCATCTCCGGCCTTAACCTAACGTCTCCGTGTTA	276
Db	391	NNTTT	450
QY	277	CTCTGAAGTTTCTGCCTTAGAGCCTCGATCGCTCACCGCATGCAATCTGTCTCGA	336
Db	451	TTTTNNTT	510
QY	337	TTTCTCTTTTCTCGCGTGAAAAATGCCCTAAATGTTCTCGATTTGGAAGTTTGTG	396
Db	511	TT	570
QY	397	CATAGSGTACTTTTTCCTAATAATTATAGTCTTAGGTAACGATACGCGCTTAC	456
Db	571	TT	630
QY	457	TGTTTTGTCATTGTTGTTGCTTTCACGCTTAGTCGCGATACGAGATATGACTGT	516
Db	631	TT	690
QY	517	GAAAAATCCCTGGTTTTTGTTTTGTTTCATATAACGATGATCTACCTTTTGG	576

[illegible]

1	1116:	contig of 1116 bp	in length
1117	1216:	gap of 100 bp	
1217	2236:	contig of 1120 bp	in length
2237	2436:	gap of 100 bp	
2437	3561:	contig of 1125 bp	in length
3562	3661:	gap of 100 bp	
3662	4781:	contig of 1120 bp	in length
4782	4881:	gap of 100 bp	
4882	6001:	contig of 1120 bp	in length
6002	6101:	gap of 100 bp	
6102	7244:	contig of 1143 bp	in length
7245	7344:	gap of 100 bp	
7345	8458:	contig of 1114 bp	in length
8459	8558:	gap of 100 bp	
8559	9640:	contig of 1082 bp	in length
9641	9740:	gap of 100 bp	
9741	10846:	contig of 1106 bp	in length
10847	10946:	gap of 100 bp	
10947	12103:	contig of 1157 bp	in length
12104	12203:	gap of 100 bp	
12204	13300:	contig of 1117 bp	in length
13321	13420:	gap of 100 bp	
13421	14520:	contig of 1100 bp	in length
14521	14620:	gap of 100 bp	
14621	15773:	contig of 1153 bp	in length
15774	15873:	gap of 100 bp	
15874	16991:	contig of 1118 bp	in length
16992	17091:	gap of 100 bp	
17092	18340:	contig of 1149 bp	in length
18241	18340:	gap of 100 bp	
18341	19476:	contig of 1136 bp	in length
19477	19576:	gap of 100 bp	
19577	20721:	contig of 1145 bp	in length
20722	20821:	gap of 100 bp	
20822	21942:	contig of 1121 bp	in length
21943	22042:	gap of 100 bp	
22043	22141:	contig of 1099 bp	in length
22142	22341:	gap of 100 bp	
22342	24372:	contig of 1131 bp	in length
24373	24472:	gap of 100 bp	
24473	25627:	contig of 1155 bp	in length
25628	25728:	gap of 100 bp	
25728	26888:	contig of 1161 bp	in length
26889	26989:	gap of 100 bp	
26989	28030:	contig of 1042 bp	in length
28031	28130:	gap of 100 bp	
28131	29303:	contig of 1173 bp	in length
29304	29403:	gap of 100 bp	
29404	30554:	contig of 1151 bp	in length
30555	30654:	gap of 100 bp	
30655	31751:	contig of 1097 bp	in length
31752	31851:	gap of 100 bp	
31852	33966:	contig of 1115 bp	in length
33967	33066:	gap of 100 bp	
33067	34113:	contig of 1047 bp	in length
34114	34213:	gap of 100 bp	
34214	35329:	contig of 1116 bp	in length
35330	35429:	gap of 100 bp	
35430	36574:	contig of 1145 bp	in length
36575	36674:	gap of 100 bp	
36675	37828:	contig of 1154 bp	in length
37829	37928:	gap of 100 bp	
37929	39069:	contig of 1135 bp	in length
39069	39164:	gap of 100 bp	
39164	40288:	contig of 1126 bp	in length
40289	40389:	gap of 100 bp	
40390	41471:	contig of 1082 bp	in length
41472	41571:	gap of 100 bp	
41572	42634:	contig of 1063 bp	in length
42635	42734:	gap of 100 bp	
42735	43904:	contig of 1170 bp	in length
43905	44004:	gap of 100 bp	
44005	45188:	contig of 1124 bp	in length

*	451.29	452.28	gap of 100 bp	in length
*	452.29	453.58	contig of 1130 bp	in length
*	463.59	464.58	gap of 100 bp	in length
*	466.59	475.95	contig of 1137 bp	in length
*	475.96	476.95	gap of 100 bp	in length
*	476.96	487.85	contig of 1090 bp	in length
*	487.86	488.85	gap of 100 bp	in length
*	488.86	499.60	contig of 1075 bp	in length
*	499.61	506.60	gap of 100 bp	in length
*	506.61	511.32	contig of 1072 bp	in length
*	511.33	512.32	gap of 100 bp	in length
*	512.33	523.56	contig of 1124 bp	in length
*	523.57	524.56	gap of 100 bp	in length
*	524.57	535.56	contig of 1100 bp	in length
*	535.57	536.56	gap of 100 bp	in length
*	536.57	547.82	contig of 1126 bp	in length
*	547.83	548.82	gap of 100 bp	in length
*	548.83	560.08	contig of 1126 bp	in length
*	560.09	561.08	gap of 100 bp	in length
*	561.09	573.07	contig of 1199 bp	in length
*	573.08	574.07	gap of 100 bp	in length
*	574.08	585.29	contig of 1122 bp	in length
*	585.30	586.29	gap of 100 bp	in length
*	586.30	597.93	contig of 1170 bp	in length
*	598.00	598.99	gap of 100 bp	in length
*	599.00	609.83	contig of 1084 bp	in length
*	609.84	610.83	gap of 100 bp	in length
*	610.84	621.93	contig of 1110 bp	in length
*	621.94	622.93	gap of 100 bp	in length
*	622.94	633.97	contig of 1104 bp	in length
*	633.98	634.97	gap of 100 bp	in length
*	634.98	645.74	contig of 1077 bp	in length
*	645.75	646.74	gap of 100 bp	in length
*	646.75	658.00	contig of 1126 bp	in length
*	658.01	659.00	gap of 100 bp	in length
*	659.01	669.93	contig of 1093 bp	in length

ORIGIN

Query Match	10.1%;	Score 70;	DB 2;	Length 66993;
Best Local Similarity	41.2%;	Pred. No. 0.00018;		
Matches 214; Conservative	0;	Mismatches 305;	Indels 0;	Gaps 0;

QY	13	IIIIIIAAGCGTAAAGACGCTCCGATCTGCTCTCTGCGAGTAATATCTCTCC	216
Db	56776	TT	56833
QY	217	GGAGTCTGACTACTACTCTCACTCCGGGCGCTTAAACTAGCTCCGCGTTA	276
Db	56836	TTANNNTNNNTTTNTTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	56895
QY	277	CTCTGAAGTTTCTGCGTTAGAGCGCTCGACTCAGCGATGCTGTGCTGCA	336
Db	56896	TTTTNTNNNNNNNTTT	56955
QY	337	TTTCTCTTTTCTGCGTGAAGAAATGCGCTATGCTTCGATTTGGAAGTTTGTG	396
Db	56956	TT	57015
QY	397	CTAAGGTTACTTTTTCCCTATATTATTAAGTCTTAGTAAAGATACCTGCGCTTAC	456
Db	57016	TTNNNNNNNTTTN	57075
QY	457	TGTTTTTGTCAATTTGTGTGCTTACCGGTTAAGTCGTGAGAGTATTGACTGT	516

[illegible]

RESULT 10

LOCUS	AJ592058	1434 bp	DNA linear	PLN 23-OCT-2003
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence, right border, clone 602D05.			
ACCESSION	AJ592058			
VERSION	AJ592058.1	GI:37941682		
KEYWORDS	right border; T-DNA flanking sequence.			
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			

REFERENCE

AUTHORS
 Bismuth, V., Balzergue, S., Vuorecquy, B., Amoult, S., Gombou, F.,
 Chauvin, S., Bechtold, N., Cruaud, C., Derose, R., Pelletier, G.,
 Lepiniec, L., Caboche, M. and Lecharny, A.
 TITLE
 T-DNA integration into the Arabidopsis genome depends on sequences
 of pre-insertion sites
 JOURNAL
 EMBO Rep. 3 (12), 1152-1157 (2002)

REFERENCE

REFERENCE 2 (bases 1 to 1434)

JOURNAL

COMMENT

PCR was performed on DNA from transformants of *Arabidopsis thaliana* plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publicities/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiosgen.fr>).

FEATURES

```
misc_feature
1..1434
    /organism="Arabidopsis thaliana"
    /mol_type="genomic DNA"
    /culivar="Massiliaewskaja"
    /db_xref="taxon:3702"
    /clone="602D05"
    /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
    /note="T-DNA flanking sequence
right border"
```

ORIGIN

[illegible]

Db 2225 GTTATTTTGTCTTGT 2233

RESULT 12
AC022680
LOCUS AC022680
DEFINITION Homo sapiens chromosome 8 clone RP11-511E14 map 8, LOW-PASS
ACCESSION AC022680
VERSION AC022680.2 GI:9143927
KEYWORDS HTG, HTGS PHASED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE 1 (bases 1 to 88932)
JOURNAL Homo sapiens chromosome 8, clone RP11-511E14
REFERENCE
AUTHORS 2 (bases 1 to 88932)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckertly,R., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Chenopel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,W., Fensholt,U.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,T.,
Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lien,C., Karatas,A., Klein,J.,
Macdonald,P., Margulis,N., McEwan,P., McKernan,K.,
McPheters,R., Melidim,J., Meneus,L., Morrow,T., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Ollivar,T.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,S.V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Teeffaye,S., Theodore,J.,
Tirrel,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
DIRECT Submission
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6910620.
All repeats were identified using RepeatMasker:
Smtc://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WITB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5106
Center clone name: 51L_E_14

* NOTE: This record contains 86 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 933: contig of 933 bp in length
* 934 1033: gap of 100 bp
* 1034 1961: contig of 928 bp in length
* 1962 2061: gap of 100 bp
* 2062 3005: contig of 944 bp in length
* 3006 3105: gap of 100 bp
* 3106 4066: contig of 961 bp in length
* 4067 4166: gap of 100 bp
* 4167 5120: contig of 954 bp in length

*	42644:	contig of 939 bp in length
*	43583:	gap of 100 bp
*	43683:	contig of 957 bp in length
*	44640:	44739: gap of 100 bp
*	44740:	45670: contig of 931 bp in length
*	45671:	45770: gap of 100 bp
*	45771:	46715: contig of 945 bp in length
*	46716:	46815: gap of 100 bp
*	46816:	47746: contig of 931 bp in length
*	47747:	47846: gap of 100 bp
*	47847:	48772: contig of 926 bp in length
*	48773:	48872: gap of 100 bp
*	48873:	49804: contig of 932 bp in length
*	49805:	49904: gap of 100 bp
*	49905:	50849: contig of 945 bp in length
*	50850:	50949: gap of 100 bp
*	50950:	51842: contig of 893 bp in length
*	51843:	51942: gap of 100 bp
*	51943:	52882: contig of 940 bp in length
*	52883:	52982: gap of 100 bp
*	52983:	53928: contig of 946 bp in length
*	53929:	54028: gap of 100 bp
*	54029:	54988: contig of 960 bp in length
*	54989:	55088: gap of 100 bp
*	55089:	56034: contig of 946 bp in length
*	56035:	56134: gap of 100 bp
*	56135:	57063: contig of 929 bp in length
*	57064:	57163: gap of 100 bp
*	57164:	58097: contig of 934 bp in length
*	58098:	58197: gap of 100 bp
*	58198:	58135: contig of 938 bp in length
*	59136:	59235: gap of 100 bp
*	59236:	60161: contig of 926 bp in length
*	60162:	60262: gap of 100 bp
*	60262:	61180: contig of 919 bp in length
*	61181:	61280: gap of 100 bp
*	61281:	62213: contig of 933 bp in length
*	62214:	62313: gap of 100 bp
*	62314:	63241: contig of 928 bp in length
*	63242:	63341: gap of 100 bp
*	63342:	64242: contig of 901 bp in length
*	64243:	64342: gap of 100 bp
*	64343:	65301: contig of 959 bp in length
*	65302:	65401: gap of 100 bp
*	65402:	66365: contig of 964 bp in length
*	66366:	66465: gap of 100 bp
*	66466:	67381: contig of 916 bp in length
*	67382:	67481: gap of 100 bp
*	67482:	68455: contig of 974 bp in length
*	68456:	68555: gap of 100 bp
*	68556:	69479: contig of 924 bp in length
*	69480:	69579: gap of 100 bp
*	69580:	70512: contig of 933 bp in length
*	70513:	70612: gap of 100 bp
*	70613:	71530: contig of 918 bp in length
*	71531:	71630: gap of 100 bp
*	71631:	72564: contig of 934 bp in length
*	72565:	72664: gap of 100 bp
*	72665:	73601: contig of 937 bp in length

ch 9.9%; Score 69; DB 2; Length 88932;
 1 Similarity 45.0%; Pred. No. 0.0002;
 165; Conservative 0; Mismatches 202; Indels 0; Gaps 0;
 5 TTTCTGTCGATTTCTCTTTTCTTCGCGCGAAATAGCCCAATGTCGATTTG 384
 6 TTT 6504
 5 AAGTTTTTGTCTATGGGTTACTTTTTCCATAATTATAGTTCTTAGTAACATA 444
 6 TTT 6510
 5 CCTGGCTACTGTTTGTTCATTTGTGTGCTTACCGCTTAGTCGCGTAAACGA 504

[illegible]

RESULT	13			
AX346599				
LOCUS				
DEFINITION	Sequence 1670 from Patent WO0200928.	6668 bp	DNA	linear
ACCESSION	AX346599			PAT 01-FEB-2002
VERSION	AX346599.1			
KEYWORDS	GI:18494485			
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS	Olek, A., Piepenbrock, C. and Berlin, K.			
TITLE	Diagnosis of diseases associated with the immune system			
JOURNAL	Patent: WO 0200928-A 1670 03-JAN-2002;			
FEATURES				
source				
	Epigenomics AG (DE)			
	Location/Qualifiers			
	1..6668			
	/organism="synthetic construct"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:32630"			
	/note="chemically prepared genomic DNA (Homo sapiens)"			
ORIGIN				

Query Match	9.9%	Score 68.8	DB 6	Length 6668
Best Local Similarity	45.8%	Pred. No. 0.00043		
Matches 238	Conservative 0	Mismatches 282	Indels 0	Gaps 0
QY	156	GTATTAGAGCTTAATAAGCTTCATTTGCTCTCTCGTCAGTTAATTTCTCCGC	215	
Db	2742	GTATTATTGTTT	2801	
QY	216	CGAGTCTGACTACTACTCTCAGTCGCGGCGCTTAACTTACGTTCCGTCGTTT	275	
Db	2802	TT	2861	
QY	276	AATCTGTAAGTTTCTGCTTAGAGCTCCGATCGCCTACCGCATGATCTGTGCTCG	335	
Db	2862	TT	2921	
QY	336	ATTCTCTTTTCTGCTCGCGAATAAATGCGCCAAAGTCTCGATTGCAAGTTTTGT	395	
Db	2922	TT	2981	
QY	396	GCATAGGTTACTTTTCCCTAATATTTAATGTTCTTAGTACAGATACCGTGCCTTA	455	
Db	2982	TTTTTTTGTT	3041	
QY	456	CTGTTTTGTCATTGTTGTGTCCTTCAACGTTAGTCGATCGAGATTTGACTG	515	
Db	3042	TTTTTTTTTGTTGT	3101	
QY	516	TGAAAAATCCCTGCTTTTGTGTTTTGTTTCAATAAATCGAATGATCTACCTTTGT	575	

```
Db      3102 TGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGT 3161
Oy      576 GCTTGAATGTTGTTTGTAGCGCTAGCGTGTGCTGTTAATCACTCAGTTCATG 635
Db      3162 TTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3221
Oy      636 TGTGATTTTGAGATTTTGCTAGTGTGCTGTTCTTT 675
Db      3222 TTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3261

RESULT 14
AC022851      81120 bp   DNA      linear   HTG 13-JUL-2000
LOCUS      AC022851
DEFINITION      Homo sapiens chromosome 11 clone RP11-284N16 map 11, LOW-PASS
SEQUENCE SAMPLING.
AC022851
AC022851.2   GI:9158172
VERSION      AC022851.2
KEYWORDS      HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1. (bases 1 to 81120)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE      Homo sapiens chromosome 11, clone RP11-284N16
JOURNAL      Unpublished
REFERENCE      2. (bases 1 to 81120)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckert,R., Bede,F.,
            Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
            Choeel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            Dearlano,K., Dewar,K., Domini,M., Doyle,M., Fensholt,J.,
            Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heatord,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marguis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
            Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Jul 13, 2000 this sequence version replaced gi:6922196.
            All repeats were identified using RepeatMasker:
            Smit, A.P.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----- Project Information
            Center project name: L6170
            Center clone name: 284_N_16
            -----
            * NOTE: This record contains 82 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            1          910: contig of 910 bp in length
            911          1010: gap of 100 bp
            1011          1910: contig of 900 bp in length
            1911          2010: gap of 100 bp
            2011          2917: contig of 907 bp in length
            2918          3017: gap of 100 bp
            3018          3921: contig of 904 bp in length
            3922          4021: gap of 100 bp
            4022          4903: contig of 882 bp in length
            4904          5003: gap of 100 bp
            5004          5874: contig of 871 bp in length
            5875          5974: gap of 100 bp
            5976          6860: contig of 886 bp in length
            6861          6960: gap of 100 bp
            6961          7856: contig of 896 bp in length
            7857          7957: gap of 100 bp
            7958          8860: contig of 904 bp in length
            8861          8960: gap of 100 bp
            8961          9822: contig of 862 bp in length
            9823          9922: gap of 100 bp
            9923          10818: contig of 896 bp in length
            10819          10918: gap of 100 bp
            10919          11817: contig of 899 bp in length
            11818          11917: gap of 100 bp
            11917          12797: contig of 880 bp in length
            12798          12897: gap of 100 bp
            12898          13776: contig of 879 bp in length
            13777          13876: gap of 100 bp
            13877          14756: contig of 880 bp in length
            14757          14856: gap of 100 bp
            14857          15788: contig of 882 bp in length
            15789          15838: gap of 100 bp
            15839          16723: contig of 885 bp in length
            16724          16823: gap of 100 bp
            16824          17699: contig of 876 bp in length
            17700          17799: gap of 100 bp
            17800          18699: contig of 900 bp in length
            18700          18799: gap of 100 bp
            18800          19689: contig of 890 bp in length
            19690          19789: gap of 100 bp
            19790          20687: contig of 908 bp in length
            20688          20797: gap of 100 bp
            20798          21638: contig of 841 bp in length
            21639          21738: gap of 100 bp
            21739          22633: contig of 895 bp in length
            22634          22733: gap of 100 bp
            22734          23621: contig of 888 bp in length
            23622          23721: gap of 100 bp
            23722          24617: contig of 896 bp in length
            24618          24717: gap of 100 bp
            24718          25612: contig of 895 bp in length
            25613          25712: gap of 100 bp
            25713          26579: contig of 867 bp in length
            26580          26679: gap of 100 bp
            26680          27577: contig of 898 bp in length
            27578          27677: gap of 100 bp
            27679          28569: contig of 892 bp in length
            28570          28679: gap of 100 bp
            28680          29681: contig of 912 bp in length
            29682          29681: gap of 100 bp
            29683          30578: contig of 897 bp in length
            30579          30678: gap of 100 bp
            30679          31580: contig of 902 bp in length
            31581          31680: gap of 100 bp
            31681          32568: contig of 888 bp in length
            32569          32668: gap of 100 bp
            32670          33562: contig of 894 bp in length
            33563          33662: gap of 100 bp
            33664          34573: contig of 911 bp in length
            34574          34673: gap of 100 bp
            34675          35579: contig of 906 bp in length
            35580          35679: gap of 100 bp
            35680          36547: contig of 868 bp in length
            36548          36647: gap of 100 bp
```

36648	37546:	cont: of 899	bp	in length
37547	37646:	gap of 100	bp	in length
37647	38553:	cont: of 907	bp	in length
38554	38653:	gap of 100	bp	in length
38654	39526:	cont: of 873	bp	in length
39527	39626:	gap of 100	bp	in length
39627	40521:	cont: of 895	bp	in length
40522	40621:	gap of 100	bp	in length
40622	41509:	cont: of 888	bp	in length
41510	41609:	gap of 100	bp	in length
41610	42520:	cont: of 911	bp	in length
42521	42620:	gap of 100	bp	in length
42621	43507:	cont: of 887	bp	in length
43508	43607:	gap of 100	bp	in length
43608	44234:	cont: of 627	bp	in length
44235	44334:	gap of 100	bp	in length
44335	45212:	cont: of 878	bp	in length
45213	45312:	gap of 100	bp	in length
45313	46197:	cont: of 885	bp	in length
46198	46297:	gap of 100	bp	in length
46298	47174:	cont: of 877	bp	in length
47175	47274:	gap of 100	bp	in length
47275	48171:	cont: of 897	bp	in length
48172	48271:	gap of 100	bp	in length
48272	49157:	cont: of 886	bp	in length
49158	49257:	gap of 100	bp	in length
49258	50125:	cont: of 895	bp	in length
50153	50252:	gap of 100	bp	in length
50253	51155:	cont: of 903	bp	in length
51156	51255:	gap of 100	bp	in length
51256	52145:	cont: of 890	bp	in length
52146	52245:	gap of 100	bp	in length
52246	53156:	cont: of 911	bp	in length
53157	53255:	gap of 100	bp	in length
53257	54139:	cont: of 883	bp	in length
54140	54239:	gap of 100	bp	in length
54240	55149:	cont: of 910	bp	in length
55150	55249:	gap of 100	bp	in length
55250	56130:	cont: of 881	bp	in length
56131	56230:	gap of 100	bp	in length
56231	57112:	cont: of 862	bp	in length
57113	57212:	gap of 100	bp	in length
57213	58095:	cont: of 883	bp	in length
58096	58195:	gap of 100	bp	in length
58196	59098:	cont: of 903	bp	in length
59099	59198:	gap of 100	bp	in length
59199	60093:	cont: of 895	bp	in length
60094	60193:	gap of 100	bp	in length
60194	61096:	cont: of 903	bp	in length
61097	61196:	gap of 100	bp	in length
61197	62069:	cont: of 873	bp	in length
62070	62169:	gap of 100	bp	in length
62170	63036:	cont: of 867	bp	in length
63037	63136:	gap of 100	bp	in length
63137	64048:	cont: of 912	bp	in length
64049	64148:	gap of 100	bp	in length
64149	65036:	cont: of 888	bp	in length
65037	65136:	gap of 100	bp	in length
65137	66020:	cont: of 884	bp	in length
66021	66120:	gap of 100	bp	in length
66121	66984:	cont: of 864	bp	in length
66985	67084:	gap of 100	bp	in length
67085	68014:	cont: of 930	bp	in length
68015	68114:	gap of 100	bp	in length
68115	69032:	cont: of 918	bp	in length
69033	69132:	gap of 100	bp	in length
69133	70018:	cont: of 886	bp	in length

Db	961	TTTTTTTNNNNNTTTTTTTTTNNNNTTTTTTTTTTTTTNNNTTTTTTTTTT	9330
OY	235	TCTCAGTCGCGGCTTTAACTTAGCTCCGTGGTTACTCTGAAGTTTCGCC	294
Db	9321	TTTNT	9380
OY	295	TTAAGCCTCGATCGCCTCACGCATCATCTGCTCGATTCTCTTTCTCGCT	354
Db	9381	TTTTTTTTTTTTTTTTTTTTTTTTNNNTTTTTTTTTTTTTTTTTTTTTTTT	9440
OY	355	GAAAAATGCCCCATAGTCTCGATTCGAAGTTTGCTATGGATTACTTTTC	414
Db	9441	TT	9500
OY	415	CCATAATTAAAGTCTTAGAGACATACCCTGCGCTACTGTTTTGTCAATTTGT	474
Db	9501	TTTTTTTTNNNNNTTTTTTTTTNNNTTTTTTTTTTTTTTTTTTTTTTTTT	9560
OY	475	TGTCCTTCAACGGTTAGTGCATGACGAGATTAAGTGA AAAATCCTTCGTTTT	534
Db	9561	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNNNNNNNNNTTTTTT	9620
OY	535	TGTTTTTGTTTATATAATCGAGATGATCTAACCTTTGTGCTTGAATGTTGTTT	594
Db	9621	TTTTNNNTTTTTTTTTTTTTTTTTTTTTTTTTNNNNNTTTTTTTTGTTTTTTT	9680
OY	595	GAGCCATAGCGTTGTGAGCTGTATAACTACGTTCAATGTGGAATTTGAAATTTG	654
Db	9681	TTTTTTTTNNNTGTTTTTTTTNGTTTTTTGTGNTGNNTTGTNNNTGTGNGGTTTG	9740
OY	655	GTAGTGA CTGGGTTCTTTG TG 679	
Db	9741	TGTGGGGGCTTGTTTTGGGCTGGGG 9765	

	RESULT	15
AJ592180	LOCUS	
AJ592180	DEFINITION	810 bp DNA linear PLN 23-OCT-2003 Arabidopsis thaliana T-DNA flanking sequence, right border, clone 60EH03.
	ACCESSION	AJ592180
	VERSION	AJ592180.1 GI:37941804
	KEYWORDS	right border; T-DNA flanking sequence.
	SOURCE	Arabidopsis thaliana (chale crese)
	ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi-
REFERENCE		
AUTHORS		Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F., Chauvin,S., Bechtold,N., Craud,C., Desobe,R., Pelleter,G., Lepiniec,L., Caboche,M. and Lecharny,A. T-RNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites EMBO Rep. 3 (12), 1152-1157 (2002) MEDLINE 22363535 PUBMED 12446565 REFERENCES 2 (bases 1 to 810) AUTHORS Balzergue,S. TITLE Direct Submission COMMENT Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at http://dbsgap Versailles.inra.fr/publiclines/. This sequence has been generated in the framework of the French plant genomics program "Genoplante" (http://www.genoplante.com) and

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 19:34:49 ; Search time 165.206 Seconds
(without alignments)
11001.196 Million cell updates/sec

Title: US-10-030-829-1
Perfect score: 3275
Sequence: 1 gcaacaacaacaataa.....caaatataggtacacaaat 3275

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85.6	2.6	7218	1	US-08-232-463-14 Sequence 14, Appl
2	74.4	2.3	272	4	US-09-313-294A-121 Sequence 121, App
3	72.2	2.2	274	4	US-09-313-294A-463 Sequence 463, App
4	68.6	2.1	7218	1	US-08-232-463-14 Sequence 14, Appl
5	52.6	1.6	7286	3	US-09-331-581-3 Sequence 3, Appl
6	52.6	1.6	7938	3	US-09-331-581-14 Sequence 14, Appl
7	50.6	1.5	396	4	US-09-640-173-53 Sequence 53, Appl
8	50.6	1.5	396	4	US-09-640-173-53 Sequence 53, Appl
9	50.6	1.5	10619	4	US-10-204-708-4 Sequence 4, Appl
10	50.2	1.5	5562	4	US-10-204-708-63 Sequence 63, Appl
11	49.4	1.5	19124	2	US-08-487-826B-13 Sequence 13, Appl
12	49.2	1.5	6040	4	US-10-204-708-69 Sequence 69, Appl
13	48.6	1.5	51259	3	US-08-781-891-209 Sequence 209, App
14	48.6	1.5	51259	4	US-09-618-166-209 Sequence 14, App
15	47.8	1.5	2447	2	US-09-014-969-14 Sequence 29, Appl
16	47.4	1.4	5666	4	US-10-204-708-29 Sequence 29, Appl
17	47.4	1.4	8607	4	US-10-204-708-72 Sequence 72, Appl
18	47.4	1.4	19233	4	US-10-204-708-45 Sequence 45, Appl
19	47.2	1.4	8961	4	US-10-204-708-79 Sequence 79, Appl
20	47.2	1.4	6617	4	US-09-976-594-268 Sequence 268, App
21	46	1.4	3489	2	US-08-728-323A-1 Sequence 1, Appl
22	46	1.4	3489	4	US-09-298-568-1 Sequence 1, Appl
23	46	1.4	3489	4	US-09-410-399-1 Sequence 1, Appl
24	46	1.4	32207	2	US-08-770-379-20 Sequence 20, Appl
25	46	1.4	32207	3	US-08-757-669A-20 Sequence 20, Appl
26	46	1.4	32207	4	US-09-230-371A-20 Sequence 20, Appl
27	45.6	1.4	289	3	US-09-007-005-17 Sequence 17, Appl

C 28	45.6	1.4	289	3	US-09-244-796-17	Sequence 17, Appl
C 29	45.4	1.4	740	3	US-08-998-416-563	Sequence 563, Appl
C 30	45.4	1.4	11049	4	US-10-204-708-23	Sequence 23, Appl
C 31	44.4	1.4	9347	4	US-10-204-708-36	Sequence 36, Appl
C 32	44.4	1.4	16442	3	US-08-781-891-208	Sequence 208, App
C 33	44.4	1.4	16442	4	US-09-618-166-208	Sequence 208, App
C 34	44.2	1.3	1493	1	US-08-340-820-24	Sequence 24, Appl
C 35	44.2	1.3	1493	1	US-08-593-535-24	Sequence 24, Appl
C 36	44	1.3	240	1	US-08-628-417-6	Sequence 6, Appl
C 37	44	1.3	2394	4	US-09-800-729-33	Sequence 33, Appl
C 38	44	1.3	11050	4	US-10-204-708-85	Sequence 85, Appl
C 39	44	1.3	19233	4	US-10-204-708-46	Sequence 46, Appl
C 40	43.6	1.3	359	4	US-09-621-976-16008	Sequence 16008, A
C 41	43.6	1.3	2915	4	US-09-336-115C-5	Sequence 5, Appl
C 42	43.6	1.3	5501	4	US-10-204-708-38	Sequence 38, Appl
C 43	43.6	1.3	11015	4	US-10-204-708-55	Sequence 55, Appl
C 44	43.4	1.3	140	1	US-08-628-417-5	Sequence 5, Appl
C 45	43.4	1.3	359	4	US-09-621-976-16019	Sequence 16019, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHRIFFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PT99PT-F15
; US-08-232-463-14
Query Match 2.6%; Score 85.6; DB 1; Length 7218;

[illegible]

RESULT 2
US-09-313-294A-121
US-09-313-294A-121

```

/ Sequence 121, Application US/0931294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Lalgudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherran, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EARS
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 121
/ LENGTH: 272
/
/ TYPE: DNA
/
/ ORGANISM: Zea mays
/
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700548569H1
/ NAME/KEY: unsure
/ LOCATION: 14, 32, 61, 127, 162, 192
/ OTHER INFORMATION: a, t, c, g, or other
/
US-09-313-294A-121

```

Query Match	2.3%	Score 74.4;	DB 4;	Length 272;
Best Local Similarity	58.7%;	Pred. No. 2.4e-10;		
Matches 142;	Conservative	0;	Mismatches 97;	Indels 3;
				Gaps 1

Qy	1777	TATGAGCTCTTAGAGACAGCCATTCCTATGTGTCCACAGGGCCATCTGGGATGATGCTT	1836
Db	2	TATGAGCAAGTAAAGACGTCAATGCTATGTCCTCTGTGGCACCGTGGTATGAGCGTN	61
Qy	1837	CTGATGTTTGAAGCAATGCCATCTGGCTATTTTGAAGCCGAAGCCTCCACCGGAGTTA	1896
Db	62	TTATATTTTGAAGCTCAGCTGTGGGCTACATGAAAGCTGAACTTGTGATTAACACTTT	121
Qy	1897	GCTGATGAGGCTTAGATAGATTTGCTGG---GGTCAGAGCGCAGTATGTTTCTGGA	1955
Db	122	CTTAAACAAGTACAGACGAATTTATGACACCTACGCGCAGGTTGATTTGTGCTGTGT	181

QY	1954	GGGTGTCCCACTGATGATCGTTCCCTGCACACAAAGATCTGAGCAATTCATCA	2013
Db	182	GGGAAAGGACCTATATGTTCTTGCAAAAGAGATGTGACCGCATTTAAACAG	241
QY	2014	CA	2015
Db	242	CA	243

RESULT 3
US-09-313-294A-463
Sequence 463 Application US/09313294A

```

/ Sequence: 700549278H1
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Lalgudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 463
/ LENGTH: 274
/ TYPR: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incycle ID No. 6476212 700549278H1
/ NAME/KEY: unsure
/ LOCATION: 15, 53
/ OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-463

```

Query Match	2.2%	Score 72.2;	DB 4;	Length 274;
Best Local Similarity	60.6%;	Pred. No. 9, 6e-10;		
Matches 152;	Conservative	0;	Mismatches 95;	Indels 4; Gaps 2.

QY	1777	TATGAGGCTCTTAAGGACAGCCCATTCCTAATGCT--CAACAGGGCCATGCTGGAGATGACTGT	1835
Db	2	TATGAGGAGGATTAAGGACAGTATGCTCTAATGCTCCTCTGGGGGACAGCTGATGATGAGCGT	61
QY	1836	TCGTATGTTTGGAGCAGTGCACATGCTATTTTGGAGGGCCGACAGCCTCCACCGGGAGTT	1895
Db	62	GTTAATATTTGAAAGCTCAGCTGTGGGCTACATGGAACTAAAGCTGCTGCAATAAACACTT	121
QY	1896	AGCTGAGATGGGGGTTAATGATAGAAATTTGCCGG--GGTCAGAAAGCGCAAGTATGTTTTCTGG	1952
Db	122	TGTTAATCAAGGTACAGACAGGAAATTCATGTGCACTAAGCAAGTTTCATTTTGCCCTGG	181
QY	1953	AGGTTGTTCGCCAATCTGATGCGCTTCTTGCAACGAAGCAAAATCTGGCAATATTCATCA	2012
Db	182	TGGGAAAGGCAACTATATGTTTTCTTAGCAAAACAAAGGAGTATGAGGCAATTTACAA	241
QY	2013	ACACTCTCAAG	2023
Db	242	GCATTGCCAGG	252

RESULT 4
US-08-232-463-14/C

```

? Sequence 14, Application US/0622705
? Patent No. 5670367
?
? GENERAL INFORMATION:
?
? APPLICANT: DORNER, F.
? APPLICANT: SCHEIFLINGER, F.
? APPLICANT: FALKNER, F. G.
? TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
?
? NUMBER OF SEQUENCES: 52
?
? CORRESPONDENCE ADDRESS:
?
? ADDRESSEE: Foley & Lardner
?

```


Query Match	1.6%	Score 52.6;	DB 3;	Length 7938;
Best Local Similarity	56.7%;	Pred. No. 0.0013;		
Matches 97;	Conservative 0;	Mismatches 74;	Indels 0;	Gaps 0;

Qy	9	AAACAAAAATTAGAGAGATGTTCTGTTAGCAATTAATTAATAGTGGGAAACAATTAAGTTA	68
Db	5211	AACATTCATGAGAAAGATGTTCTGTGAGCTGGAATATTAATGACGTACACAGACAAAA	5270
Qy	69	AGCGAAAAAGGAAAAAAGGTACAAATAATGAAAAACAATACTGAATGAAATTT	128
Db	5271	AAAG	5330
Qy	129	GGAGTCGAGATCGAAAAACGAGCCGTTTAACTTAATAAGCTTCT	179
Db	5331	TACCTTTCGAGCGGGAAGAACGAGCCGATCCAGACATGATTAAGATCAT	5381

RESULT 7

```

US-09-640-173-53
? Sequence-53, Application US/09640173
? Patent No. 6613515
? GENERAL INFORMATION:
? APPLICANT: Xu, Jiangchun
? TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
? TITLE OF INVENTION: METHODS OF USE THEREFOR
? FILE REFERENCE: 210121.484C2
? CURRENT APPLICATION NUMBER: US/09/640,173
? CURRENT FILING DATE: 2000-08-15
? NUMBER OF SEQ ID NOS: 196
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 53
? LENGTH: 396
? TYPE: DNA
? ORGANISM: Homo sapien
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (1)...(396)
? OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

```

Query Match	1.5%;	Score 50.6;	DB 4;	Length 396;
Best Local Similarity	49.8%;	Pred. No. 0.00079;		
Matches 119; Conservative	0;	Mismatches 120;	Indels 0;	Gaps 0

Oy	TTTTTGTGCAAGGGTTACTTTTTCCCAATTTATAGTCCTAGTAACAGATCCCG	448
Db		
Oy	18 TT	77
Oy	449 CGCTTACGTATTGTTGCATTTGTGTCCTTACCGCTTAGTCGTATCGAGTAA	508
Db		
Oy	78 TT	137
Oy	509 TTGACTGTGAATAATCCTCGTTTTTGCGTTTGTTCATATAAATCGATGATCAAC	568
Db		
Oy	138 TT	197
Oy	569 CTTTGTGCTGATGTTGTTTTTGACCAGATGCGTGTGCTGTATTAACCTCA	627
Db		
Oy	198 TTTTTTTTTTTTTTTTTTTTAAATNTTTTTTTTNNCTTNNTTTTAATCA	256

```

RESULT 8
US-09-713-550-53
; Sequence 53, Application US/09713550
; Patent No. 6617109
;
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stoll, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.4844

```

CURRENT APPLICATION NUMBER: US/09/713,550
CURRENT FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 205
SOFTWARE: FastSeq for Windows Version 3.4
SEQ ID NO 53

Query Match	1.5%;	Score 50.6;	DB 4;	Length 396;
Best Local Similarity	49.8%;	Pred. No. 0.00079;		
Matches 119;	Conservative	0;	Mismatches 120;	Indels 0;
			Gaps	0;

QY		389	TTTTTGCTAAGGGTACTTTTCCTCAATATTAAAGTCTTAGGTAACGAATACCTG	448
Dd		18	TT	77
QY		449	CCHCTACGTITTTTGTCATTTGTGTNGCTTCACC GTTAGCGCATGCAGAGTAT	508
Dd		78	TT	137
QY		509	TTGACTGTGAAAATCCTTCGTTTTTTGGTTTTTGTTTCATAATAAACGATTCATCAC	568
Dd		138	TT	197
QY		569	CTTTGSGCTTGGATGTTGTTTTTGAAGCCATGCGTGTGGCTGTATAACTCCA	627
Dd		198	TTTTTTTTTTTTTTTTTTTTTTTANNNTTTTTTTTNCTTNNTTTAAATCCA	256

RESULT
US-10-2

```

US-10-204-708-4
; Sequence 4, Application US/10204708
; Patent No. 6677731
;
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
;
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
;
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
;
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
;
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
;
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
;
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
;
; SEQ ID NO 4
; SEQ ID NOS: 98
;
; LENGTH: 10619
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
;
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;
US-10-204-708-4

```

Query Match	1.5*	Score 50.6	DB 4	Length 10619
Best Local Similarity	49.8*	Pred. No. 0.0055		
Matches 128	Conservative	0	Mismatches 129	Indels 0
			Gaps 0	
404	TTACTTTTTCOCATATTTATATAGTTCTTGTATACATACGCGCTTACTGTTTT	463		

Db 15736 TTTATTAATATATTTTTTCTTTTTTTTGTGTTTAAAGATATATATTTTTTTTTT 15677
QY 617 TTTAACTTCAGCTGATGCGGATTTGAGATTTTGAGAGACTGGGTTCTT 675
Db 15676 AATGTTTTTTTTTCTTTTGTGTTTATTTTTTAAATCATTTTTTTTTATAT 15618

RESULT 12

US-10-204-708-69
; Sequence 69, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204, 708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 69
; LENGTH: 6040
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically created genomic DNA (Homo sapiens)
US-10-204-708-69

Query Match 1.5%; Score 49.2; DB 4; Length 6040;
Best Local Similarity 49.5%; Pred. No. 0.0094;
Matches 186; Conservative 0; Mismatches 183; Indels 7; Gaps 2;

QY 379 ATTTCAGAGTTTGTGCTAGGAGTTACTTTTCCATATTTTATAGTCTTAGTA 438
Db 3338 ATTTCAGAGTTTGTGCTAGGAGTTACTTTTCCATATTTTATAGTCTTAGTA 438
QY 439 ACGATACCTGCGTCTACTGTTTGTTCATTTTGTGCTTCAACGTTTAGTGC 498
Db 3398 ATAAATATATAGTTTATTTTATTTTATAGTATGTTGTTTATCGATAATAGTATG 3457
QY 499 A----TCGAGATTTGACGCTGAAAAATCTTCGTTTGTGTTTGTTCATATAA 554
Db 3458 AAGATTTCCGGGATTTGGATATTAATAAGTTGATTTGATTTATTAATTATTAAG 3517
QY 555 TCGAGTTGATCTACCTTTGTGCTTGTGATTTGTTTGGTGAACCTATGCGTTGGCT 614
Db 3518 TGGATTTGAGGTAAGTTAGTTTGTGTTTATGTTTATTTT---TATTTGTAATAATTT 3574
QY 615 TGTATTAATCTTACGTTTCATGCTGATTTTGTGATTTTGTGATGACTGTGGTTCTT 674
Db 3575 ATTAGAAGTTGTTGTAATTTGATTTTATGTTTATTAATGATTTTGTGAAAAAG 3634
QY 675 TGTGCTATAGTTGTAATAAATAGTTCTAGGCGTGTGCAATGCTAAGAAAAAGAC 734
Db 3635 TTGCTATATAGACGCTGATTAATAAATGAAGATTTGTTATTTTGAATAAATAGTAT 3694
QY 735 GTTCAGGCTGTATA 750
Db 3695 GATTAAGTATGTTATA 3710

RESULT 13

US-08-781-891-209/c
; Sequence 209, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620emburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 209:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5125 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-209

Query Match 1.5%; Score 48.6; DB 3; Length 5125;
Best Local Similarity 52.2%; Pred. No. 0.048;
Matches 108; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 2764 GAGTTTCAAGAGAAAGATGAGAGTTTGTGAAGAGAGGAGATGCTGTAATAAGT 2823
Db 232 GAGGAG 173
QY 2824 CAAG 2883
Db 172 GAGCAG 113
QY 2884 AAGAATTGATGAGGCTTTGGAACGCTCATGTCAAGCATGCGCTTCAATGAGAT 2943
Db 112 AAGGAGGAG 53
QY 2944 GATTGAGCAAAAGTCTGTACACAAG 2970
Db 52 GAGGAG 26

RESULT 14

US-09-618-166-209/c
; Sequence 209, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.

schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052,419C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 209:
US-09-618-166-209
Query Match 1.5%; Score 48.6; DB 4; Length 51259;
Best Local Similarity 52.2%; Pred. No. 0.048;
Matches 108; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 2764 GAGTTTCAGAGAAAGATGAGAGATTGTGGAAGAGAGAGATGCTGATTAAGAT 2823
DB 2332 GAGGAGGAG 173
QY 2824 CAGATTTGATCTGAG 2883
DB 172 GAGCAG 113
QY 2884 AAGAAATTTGATGAGGCTTTGGAACAGCTCATGTATCAAGCATGAGCTTCAATGAAGAT 2943
DB 112 AAGGAGGCGGAG 53
QY 2944 GATTGAGACAAAAGTCTGTACCAAG 2970
DB 52 GAGGAG 26
RESULT 15
US-09-014-969-14/c
Sequence 14, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Treacy, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2447 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-014-969-14
Query Match 1.5%; Score 47.8; DB 2; Length 2447;
Best Local Similarity 47.2%; Pred. No. 0.013;
Matches 142; Conservative 1; Mismatches 158; Indels 0; Gaps 0;
QY 390 TTTTGTGCTATGGGTACTTTTCCCTAATTTTATAGTTCTAGTAACGATACCTGC 449
DB 2447 TTTTGTGCTATGGGTACTTTTCCCTAATTTTATAGTTCTAGTAACGATACCTGC 2488
QY 450 GTCTTACGTTTGTGCTATTTGTGTGCTTACCCGTTTGTGCTGATGAGATAT 509
DB 2387 TTTTGTGCTATGGGTACTTTTCCCTAATTTTATAGTTCTAGTAACGATACCTGC 2328
QY 510 TGACGTGAAATACTGCTTTTGTGCTTATGCTTATGCTTATGCTTATGCTTATGCT 569
DB 2327 TTTTGTGCTATGGGTACTTTTCCCTAATTTTATAGTTCTAGTAACGATACCTGC 2268
QY 570 TTTTGTGCTATGGGTACTTTTCCCTAATTTTATAGTTCTAGTAACGATACCTGC 629
DB 2267 TTTTGTGCTATGGGTACTTTTCCCTAATTTTATAGTTCTAGTAACGATACCTGC 2208
QY 630 TTGATGTGAGATTTTGAATTTTGTGATGACTGTGGGTTCTTTGGTGCTATGATT 689
DB 2207 ATCAATATACAGATCCAGATATGTGAAACATATATATATATATATATATATAT 2148
QY 690 G 690
DB 2147 G 2147

Search completed: April 8, 2004, 16:29:38
Job time : 170.206 secs

This Page Blank (uspto)


```

QY 1269 CAAAGAGCCTGATCAAGAGCAGATAAGTGTCAAAAAGTTCTTGGACGTTG 1328
DB 2033 CAAAAGAACAGAGACTGTAAAGAGCAATGTGTTAAAGATTTCTTGAAGACTTG 1974
QY 1329 GATAGCTTGTGATGAGAGATTAATGACCAAGAGGAGGAGGATTTGCTCAGCTTGT 1388
DB 1973 GATGGCTTATCTATCGAAAGATCAATGAAACCGGAAAGGAGGATGCACTGATACGCTTGT 1914
QY 1389 CAGAACGACCTGTGTCATGATTTGTATA--ACTGACCTCTTACTAGCTCATGCG 1445
DB 1913 CGAGGTGTCTCCGTCTATTTGATTTGTACAGAGGCTGACGCTCTGATTAATCTATCC 1854
QY 1446 AGGACAAAAGAGCTTAGGCGAGTTAGCTCCATAGAGATTTGCTGAGATTTAGAAAAG 1505
DB 1853 AAAACAAAGGGGTCAAAAAGGGTGAAGATCCATAGGAGATTTGCTGAGCTTTTGGAAAG 1794
QY 1506 GATCTACAGATGAGAGGCGCATCTGTCAATTCCTTGTGTGAGATTTATGGGCACTGGAG 1565
DB 1793 GAATCTGCGAGAAAGGCTCTGCAATTAATCCACTGGGGAAGTATTTGTATGAGGAA 1734
QY 1566 GATTTGGGTAGAGATGAAGATTAATGAAATTTGCTGCGCTCCATAGCTCATCATG 1625
DB 1733 GGTTTAAAGATGAGAAAGATCATGAATAGTTTGGCTCGAAGTTTCTCATTTGAG 1674
QY 1626 AATCTAGACTGTATGAGAGCATTAAGGATTAAGGATTTCTTCTTTTACTTCTT 1685
DB 1673 AATCAAGGCTTGAACAGATGAATAATGA-- 1643
QY 1686 TAATTTTCTTCTGATTTCTAATGATTTTCAATTTGATGCTGGCGCATGGG 1745
DB 1642 -----AGTGGCTAGGATAGG 1627
QY 1746 CAACCAAGAGCTCTGGAATACTTCAGCAAGTATGAGGCTCTTGAAGCAGCCATTTCTTA 1805
DB 1626 TAACCAAGAGCTCTTCTTAATTAATTTTGAAGCATTAATGCTGTGAAGCAGCTTA 1567
QY 1806 TGGTCACAGGAGCATCTGTGGATGAGTTCATGATTTTGAAGCAGTGCCTGCTTA 1865
DB 1566 TGGCCCCAGGCTATCCGGGGTTGAGGTTTGAATTTGAACATCAAGTATAGGTTA 1507
QY 1866 TTTGAGGCGGAACGCTCTCCACCGGAGTTGCTGAGATGGGGTTAGATAGATTTGCTG 1925
DB 1506 TCTTGGGCTGAGCGCTTACACAGACATTTGCAAGAACAGGAACTGACGAGATGCTG 1447
QY 1926 GGGTCAGAGCGCATGATGTTT-----CTGAGGTGTTGCCAATGATGCTTCT 1979
DB 1446 GTTTAGTACACACCGATATTTCTCCCTGTGGCGGCGAGCTCTATGATACAT 1387
QY 1980 TGCACGAGAGCATGCTGAGCATATTCATCAATCAACCTCTCAAGGT 2025
DB 1386 GGCATATAAAGAGACCTGACCTTTTCAACAGACATTCGCAAGGT 1341

```

RESULT 2
US-10-424-599-91177

Sequence 91177, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 91177
TYPE: DNA
ORGANISM: Glycine max

```

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_53343C.1
US-10-424-599-91177
Query Match
Best Local Similarity 7.7%; Score 253; DB 12; Length 1377;
Matches 428; Conservative 0; Mismatches 220; Indels 13; Gaps 3;
QY 1047 TATGATTAACAATTTGTGTGCAACCCCACTGATCTCCCTCTTTGGAAGAGATGG 1106
DB 29 TATGAAGACATATGTGATCTCAATCTGATTTCCGCTCCATGACATGAGCTGG 88
QY 1107 AATTGGACGCAAGAGAGGTTCTGCTGACACACAGCTGTGACAGATTTCTTGAAGT 1166
DB 89 AACTGGCAATCCAGACCTGTGCAATTCAGTCCAAATGTAAGGATGAATTTACCAAG 148
QY 1167 G-----AGATGATGTGATTAATGCTTCTGAGGAAGAGATATCCAGCTTTG 1217
DB 149 GACCTCAAAAAGAAATTAATGATGTTGATGATGAGGAGAGAGAGAAATCTGATTTG 208
QY 1218 GATGATTCGATGACGACCTTGCAAGTATGATTAATGATGATGATGATGATGATGATG 1277
DB 209 GAAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
QY 1278 CATGATCAAGAAAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1337
DB 269 CATGACACGTAAGAAAGCAATGATGATGATGATGATGATGATGATGATGATGATG 328
QY 1338 TCGATCGACGATTAATGAACCAAGAGGCACTGATGATGATGATGATGATGATGATG 1397
DB 329 ACTGTTGAACAGATTAACCAAGAGGCACTGATGATGATGATGATGATGATGATGATG 388
QY 1398 CCGTGTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1454
DB 389 CCGTGTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 448
QY 1455 GAGCTAGGCAAGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1514
DB 449 GATCAAAAAGGAGTAAATTAATGATGATGATGATGATGATGATGATGATGATGATG 508
QY 1515 ATGAGAGGCGCATCTGTGATTTCTTGTGATGATGATGATGATGATGATGATGATGATG 1574
DB 509 AAAAGGGTACTTCAATTAATTCAGCTGGGAAAGATTTGAAAGTGAAGGTTTGA 568
QY 1575 GAGATGAAGAGATTAATTAATTTGCTGCGCTCCATGATGATGATGATGATGATGATG 1634
DB 569 GAAGAG-AAAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 627
QY 1635 CTGATTAAGACGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1694
DB 628 CTGTAACAGATGAATAATGATTAAGGATTAATTAATTAATTAATTAATTAATTAAT 687
QY 1695 T 1695
DB 688 T 688

```

RESULT 3

US-10-424-599-91176
Sequence 91176, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 91176

LENGTH: 1589
 TYPE: DNA
 ORGANISM: Glycine max
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT3847_53342C.1
 US-10-424-599-91176

Query Match 6.6%; Score 215.6; DB 12; Length 1589;
 Best Local Similarity 64.6%; Pred. No. 6.9e-44; Indels 12; Gaps 2;
 Matches 358; Conservative 0; Mismatches 184

1047 TATGATTAACACTTTTGGGACCCCACTGTATCTCGCCCTCTTGGAGAGAGATGG 1106
 29 TATGAAGACACTATGTGACTCCAAATCTGTATTCGGCTCCACATCGAGATGGCTGG 88
 1107 AATTGGCAGCAAGAGAGAGTTCTGCTCAGACACAGCTGTGACAGAGTTTCTGACGTG 1166
 89 AACTGGCAATCCAGACCTGTGTCAATTCAGTCCAAATGTAGGGATGMAATTTTCCAGAG 148
 1167 G-----AGGATGATGTGATTAATGCTTCTGAGGAAGAGATGATTCGATGCTTTG 1217
 149 GACCTTCAAAAGATTAATGTGTGTGATGATGATGGAGAGAGAGATGTGATTTG 208
 1218 GATGATTTGATGACGACCTTGGCAATGATGATTAATGATCTCGATGTGATGATGATGATG 1277
 209 GAAGATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
 1278 CATGATCAAGAAAGCAATTAAGTGTCTTCTTGGAGAGCTTGTGATGATGATGATGATG 1337
 269 CATGAACACATGTAAAGAAAGCAATGTGTAAAGAGTCTTGTGATGATGATGATGATGATG 328
 1338 TCGATGACAGATTAATGAACAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 1397
 329 ACTGTGAACAGATTAAGCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 388
 1398 CCTGTGCAATCATTTGTGATTA--ACCTGCACTCTCTACTAGCTGATGATGATGATGATG 1454
 389 CCGGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 448
 1455 GAGAGTAAAGGAGATTAAGCTCCATTAAGATTTGCTGAGAGTTTGAAGAAAGATCTACAG 1514
 449 GATCAAAAAAGGAGTGAAGATTCACAGGAGCTTGTATCTTTTGTGATGATGATGATGATG 508
 1515 ATGAGAGGGGCAATCTGCAATCTTGTGTGATGATTAATGAGGAGTGAAGGATTTGGGT 1574
 509 AGAAGGGGCTCTGCAATTAATCCACTGCGAAGTATTTGTGATGATGATGATGATGATGATG 568
 1575 GAGGATGAAAAAGA 1588
 569 GTGCACAGAAAAGA 582

RESULT 4
 US-10-425-114-18351
 ; Sequence 18351, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaka, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313) B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 18351
 ; LENGTH: 2205
 ; TYPE: DNA
 ; ORGANISM: Zea mays

FEATURE:
 OTHER INFORMATION: Clone ID: LIB3079-030-A8_FLI
 US-10-425-114-18351

Query Match 6.0%; Score 198; DB 12; Length 2205;
 Best Local Similarity 55.7%; Pred. No. 2.6e-39;
 Matches 518; Conservative 0; Mismatches 330; Indels 82; Gaps 4;

1157 TCTGACGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1216
 452 TTCTGTCCCTGAGGAGGAGAAATGTGATGAGCAATATCATCATGATGATGATGATGATGATG 511
 1217 GATGATTTCTGATGACGACCTTGCAGATGATGATGATGATGATGATGATGATGATGATGATG 1276
 512 TGACATGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 571
 1277 CCATGATCAAGAAAGCAATTAAGTGTTCAAAAAGTTCTTTGAGAGCTTGTGATGATGATG 1336
 572 TTTTGACCTCGAAAAAGCAAGAGGTTCAAGAGTTCTTTGAAGTATGATGATGATGATGATG 631
 1337 GTGATGACGAGATTAATGAACCAAGAGAGAGATGATGATGATGATGATGATGATGATGATG 1396
 632 GAGTTTGAACAAATTAATGAACAACTGGCAATGATGATGATGATGATGATGATGATGATG 691
 1397 ACCGTGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1453
 692 ACTGTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 751
 1454 AGAGATGACGAGATTAATGAACCAAGAGAGAGATGATGATGATGATGATGATGATGATGATG 1513
 752 GGGTTCTAAGAGGTTAACTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 811
 1514 GATGAGAGGCGATCTGTATCTTGTGTGATGATTAATGAGAGTGAAGGATTTGGG 1573
 812 TCGAGGGGAACTTCAAGTGTACCAAGGGGTAAACATTTGGGAATGGAAGAGAGCTG-- 869
 1574 TGAGATGAAAGAGATTAATGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1633
 870 -CAAGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928
 1634 ACTGATTAAGACGATTAAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATG 1693
 929 CTGGAAAAAGATGAAGATGATA----- 951
 1694 CTTGTGATTTACATGATCTTGAATGTTACATTTGATGATGATGATGATGATGATGATGATG 1753
 952 -----AGTGAAGGGCATGGAAACCAAG 975
 1754 AGCTGTGATTAATCTCGACAGATTAAGGCTCTTAAGACAGGCACTTCTTAATGATGATGATG 1813
 976 AGCTCTTGAATTTTGGAGATTAATGAACAGTAAAGCAATGATGATGATGATGATGATGATG 1035
 1814 AGGGCCATGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1873
 1036 CTGGGACACCTGTATGAACGTTTAAATTTGAAAGCTCAGCTGTGAGGCTACATGATGAAG 1095
 1874 CCGAAGGCTCCACCGGAGATTAAGCTGATGATGATGATGATGATGATGATGATGATGATG 1932
 1096 CTGAAGCTGTGATTAACCTTTGTTAAACAAGATTAACAAGCAAGATTAATGATGATGATG 1155
 1933 --AAGGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1990
 1156 GCAAGTTGATTTGTGCTGTGAGGAAAGGCAATTAATGATGATGATGATGATGATGATGATG 1215
 1991 AAGATCTGACATTAATTAATCAACATCTCAAGGTTCTTCCCCAAAGAAATTTGATAT 2050
 1216 AGGATATGAGGACATTTAAACAAGATGATGATGATGATGATGATGATGATGATGATGATG 1275
 2051 ATGCTTTTAAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2080
 1276 GGTCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1305

RESULT 5
 US-10-424-599-73594/c
 ; Sequence 73594, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 73594
 ; LENGTH: 2890
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_37470C.1
 US-10-424-599-73594

Query Match 6.0%; Score 197.8; DB 12; Length 2890;
 Best Local Similarity 69.4%; Pred. No. 3,6e-39;
 Matches 284; Conservative 0; Mismatches 122; Indels 3; Gaps 1;
 QY 1243 GTGATGATTATGACTCGGATGTGATCAAAAGAGCCATGATCCGAAAGCAATTAAGT 1302
 DB 2890 GTGATGATTATGATTCGTATCTAGTCAAAATTAAGCCAGACTCGTAAAGAGCAAGG 2831
 QY 1303 GGTTCAAAAAGTTCTTTGGCAGCTTGATAGCTTGTCATCGACAGATTAATGAACAC 1362
 DB 2830 GGTTCAGAGTTCTTTGAGAACTTGATGGCTTATCTATCGAAAGATCAATGAACCGG 2771
 QY 1363 AGAGCAGTGGCAGTTCCTGCTGCTGTCAGAACGACCTGGTCATGATGTTGTA--- 1419
 DB 2770 AAAGCAGTGGCAGTTCCTGCTGCTGTCAGAACGACCTGGTCATGATGTTGTA--- 1419
 QY 1420 ACCTGACCTCTTACTAGCTCATGCGAGACAAAGAGGCTAGGCGGTTAACTCCATA 1479
 DB 2710 GACTGACGCTCTGATTAATCACTCAATCCAAACAAAGGGGTAAAGGGGTGAATCCATA 2651
 QY 1480 GAGAAATGGCTGAAGTTTGAAGAAAGGATTTACAGATGAGAGGCGCATCTGTCATTCCTT 1539
 DB 2650 GGAAGTTTGGCTGAGCTTTTGAAGAGGAACTGCGCAAAAGGGGCTGCGAGTATTCAC 2591
 QY 1540 GTGATGATTATTAAGGAGTGGAGGAGGTTGGGTGAGATGAAGATTAATGAATG 1599
 DB 2590 CTGGGAGAGTATTTGTAAGTGAAGGTTTAAAGATGATGAAGATCATGAATAG 2531
 QY 1600 TCTGGCTTCAATGTCATCATCATGAAATCTAGCTGATTAAGACGA 1648
 DB 2530 TTGGCTTCAATGTCATCATCATGAAATCTAGCTGATTAAGACGA 2482

RESULT 6
 US-10-424-599-73591/c
 ; Sequence 73591, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 73591
 ; LENGTH: 1375

TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(1375)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_37468C.1
 US-10-424-599-73591

Query Match 5.2%; Score 169.8; DB 12; Length 1375;
 Best Local Similarity 69.0%; Pred. No. 3e-32;
 Matches 247; Conservative 0; Mismatches 108; Indels 3; Gaps 1;
 QY 1243 GTGATGATTATGACTCGGATGTGATCAAAAGAGCCATGATCCGAAAGCAATTAAGT 1302
 DB 1375 GTGATGATTATGATTCGTATCTAGTCAAAATTAAGCCAGACTCGTAAAGAGCAAT 1316
 QY 1303 GGTTCAAAAAGTTCTTTGGCAGCTTGATAGCTTGTCATCGACAGATTAATGAACAC 1362
 DB 1315 GGTTCAGAGTTCTTTGAGAACTTGATGGCTTATCTATCGAAAGATCAATGAACCGG 1256
 QY 1363 AGAGCAGTGGCAGTTCCTGCTGCTGTCAGAACGACCTGGTCATGATGTTGTA--- 1419
 DB 1255 AAAGCAGTGGCAGTTCCTGCTGCTGTCAGAACGACCTGGTCATGATGTTGTA--- 1419
 QY 1420 ACCTGACCTCTTACTAGCTCATGCGAGACAAAGAGGCTAGGCGGTTAACTCCATA 1479
 DB 1195 GACTGACGCTCTGATTAATCACTCAATCCAAACAAAGGGGTAAAGGGGTGAATCCATA 1136
 QY 1480 GAGAAATGGCTGAAGTTTGAAGAAAGGATTTACAGATGAGAGGCGCATCTGTCATTCCTT 1539
 DB 1135 GGAAGTTTGGCTGAGCTTTTGAAGAGGAACTGCGCAAAAGGGGCTGCGAGTATTCAC 1076
 QY 1540 GTGATGATTATTAAGGAGTGGAGGAGGTTGGGTGAGATGAAGATTAATGAAT 1597
 DB 1075 CTGGGAGAGTATTTGTAAGTGAAGGTTTAAAGATGAGGAGGAGATCATGAAT 1018

RESULT 7
 US-10-424-599-36215/c
 ; Sequence 36215, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 36215
 ; LENGTH: 1103
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(1103)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_132704C.1
 US-10-424-599-36215

Query Match 3.2%; Score 103.8; DB 12; Length 1103;
 Best Local Similarity 67.3%; Pred. No. 1.7e-15;
 Matches 175; Conservative 0; Mismatches 83; Indels 2; Gaps 2;
 QY 2126 ATTCAATGAGCAAAAGAGCTGAATTTGAGTTGAATCATACAGAG-ATGTTGTA 2184
 DB 359 ATTCCAAAGTAATCTAAGCTCAATATGACATATGATCAGAGAAATGTTGTA 300

QY 2185 AAGAGCTGAGGAGATCTCTGAGACAAATCAGAGCTGAATCTTAAAGAACAGCTC 2244
DB 2299 CACCAATTCGGGAATGATGAGACCAACAGAGCTGCTATTTTAAAGAACAGCTC 240
QY 2245 TCAAAACAGAACAGACGCCAAGGCTGCTGAGGAATCTCTGAAATTTATGACGAGAG 2304
DB 2339 GTCGAAGGTCGCAAAACACAAAAGCTCTGAAAGAACTATGATTAATGCTGAGAG 180
QY 2305 CTGCGTGAAGCTGACGAGATATCGAGTGTGAGACA-GAGAACTAAGTGAAGATGA 2363
DB 179 CTGCGAAGCAATGAGGAAATCGCATTTGAGGAGGAGAACTAAATGCAACATGA 120
QY 2364 ACAGAACAGGAGAGATAT 2383
DB 119 AGAGACCAAAAAGMAGATGT 100

RESULT 8

US-10-312-841-2
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/MO
; CURRENT APPLICATION NUMBER: US/10/312, 841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match 2.5%; Score 81.4; DB 14; Length 3673778;
Best Local Similarity 45.6%; Pred. No. 2.2e-07;
Matches 286; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

QY 80 AAAAAAAAGTACAAAATGAAACAAATCAACTGATGAATGAATTTGAGTCCAGAA 139
DB 971449 AAAAAAAAGTACAAAATGAAACAAATCAACTGATGAATGAATTTGAGTCCAGAA 139
QY 140 TCGGAAAAAGAGCGGCTTTAGAGCTTAATAAGCTTCTCATTTGCTCTCTGCTCA 199
DB 971509 AAAGTAATTAGAGAAATGTTATCGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 971568
QY 200 GTTATTTTCTCTCCGAGGCTGACACTGCTGCTCCGGCGCTTAAACTT 259
DB 971569 TTTTGT 971628
QY 260 AGCTTCGCTGCTTACTCTGTAAGTTTCTGCTTAAAGCTCGAGCGCTCACCGC 319
DB 971629 TTTTGT 971688
QY 320 ATGCAATCTGCTCGATTTCTTTTCTTCCGTGAAAAAATGCGCTTAATGCTCGA 379
DB 971689 TTTTGT 971748
QY 380 TTTCGAAGTTTGTGCTATGAGGTTACTTTTCCCTAATTTTAATGTTCTTACGTA 439
DB 971749 TTTTGT 971808
QY 440 CGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 499
DB 971809 TTTTGT 971868
QY 500 TCGGAGTATTGACTGTGAAAAATCCTGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 559

DB 971869 TTTTGT 971928
QY 560 TTGATCTACTTTTGTGCTTGATGTTGTTTGTGAGCTTAAGCTTGTGCTGTA 619
DB 971929 TTTTGT 971988
QY 620 TAACCTCAGCTCATGTGAGATTTTGAATTTTGTGAGTGTGAGTGTGAGTGTGAGT 679
DB 971989 TTTTGT 972048
QY 680 GCTAATGCTTGAATAATGAGTTCTTG 706
DB 972049 AGCTTGAAGTGTAGTGTGTTATTTAG 972075

RESULT 9

US-10-424-599-99770/c
; Sequence 99770, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424, 599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 99770
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_61108C.1
US-10-424-599-99770

Query Match 2.4%; Score 77; DB 12; Length 951;
Best Local Similarity 63.0%; Pred. No. 1.1e-08;
Matches 119; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 2736 CAGAGCTGAGAGTGTCAAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGATTGT 2795
DB 606 CAAGTGTGAAATATCTTCAAGTTTCTGAGGCTCAAGCAAGAGATGAAGATTGCA 547
QY 2796 GGAAGAGAGAGATGCTGATTAAGATCAAGAGAAAGATGAGAGATGAGAGAG 2855
DB 546 GCGAGAGAGAGAGAAATTAATAATTCATGAAGAGAAAGTTGCTGAGAGAGAA 487
QY 2856 GCATCAGAGAGATTTTGTATCTGAGAGAAAGATTTGATGAGCTTTGAAACAGTCA 2915
DB 486 GCGAGTGGCAAGAGAGGTTGAACTTGAGAGAGAGATTGAGAAATGAACTGACGAGCTCAT 427
QY 2916 GTACAAAGCA 2924
DB 426 GGCAGAGTA 418

RESULT 10

US-10-311-455-1670
; Sequence 1670, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIERENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311, 455
; CURRENT FILING DATE: 2002-12-16

RESULT 12
US-10-240-453-54
Sequence 54, Application US/10240453
Publication No. US20030148326A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
TITLE OF INVENTION: Transcription
TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
TITLE OF INVENTION: with DNA Transcription
FILE REFERENCE: 5013.1009
CURRENT APPLICATION NUMBER: US/10/240,453
CURRENT FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: PCT/EP01/03973
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019056.8

```

; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 100325229.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 54
; LENGTH: 9539
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OS=10-240-453--54
```

Query March	2.0%	Score 66.8	DB 14	Length 9539
Best Local Similarity	46.7%	Pred. No. 2.1e-05		
Matches 212	Conservative	0	Mismatches 242	Indels 0
				Gaps 0

Qy	25	TTAAACTACGCTTCGCGCTGCTTACCTGTAAGATTTTTCGCGTTAGAGCGCTCGAATCG	31.0
Db	202	TTTTTTTTTTTGTAATTTTTCGTTTATTTTGTGTGTTTTTTTTTTTATTTTTTTTTTC	26.1
Qy	311	CCTCACCGAATGCAATTCGTCGATATTCCTTTTCTTCGCTGGAATAATGCCCTAA	37.0
Db	262	GTATTTTTTTTTTTTGCGATTTGTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTAAT	32.1
Qy	371	TGTTTCGATTCGAAGTTTTTGTGCTAATGGTACTTTTTTCCCTAATTTATTAAGT	43.0
Db	322	TATATGTTTTATATT	38.1
Qy	431	CTTAGTAACGATACCTCGCGCTTACTGTTTGTTCATTTGTGTGCTTCACCGTT	49.0
Db	382	TTTTTTTGTGTTTTTTTCGTTTTTTTTTTTTTTTTTTTATTTTTTTTTTGTTTTTTTTTT	44.1
Qy	491	AGTCGCTAATCGAGATTTGACTGTGAATAAATCCTCGTTTTTGGTTTTGTTTCATA	55.0
Db	442	ATT	50.1
Qy	551	TAAATCGATGATCTACCTTTGTGCTTAATGTTTGTTTTTGAGCGTAAAGCGTGT	61.0
Db	502	TTTTTTTTTTTTTTTGATT	56.1
Qy	611	GCGTTGTAAATCTCACGCTCATGCTGATTTTGATTTTGGTAGTGACTGTGGGT	67.0
Db	562	TTTTTTTTTTGTTTTATTTTTTTTTTTTTTTTATATTTTATATATTTATTTTTTTTTT	62.1
Qy	671	TCTTTGCTGCTATAGTGTGTAATAATAGTTCT	70.4
Db	622	TATATATATATATATTTTTTTTTTGAGGTAGGGTTT	65.5

```

RESULT 13
US-10-424-599-57885
: Sequence 57885, Application US/10424599
: Publication No. US20040031072A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ. ID NOS: 285684
: SEQ. ID NO 57885
: LENGTH: 673
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:

```

```

; NAME/KEY: unsure
; LOCATION: (1)..(673)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_23282C.1
US-10-424-599-57865

```

Query Match	2.0%;	Score 66.2;	DB 12;	Length 673;
Best Local Similarity	47.0%;	Pred. No. 4.8e-06;		
Matches 170; Conservative	0;	Mismatches 192;	Indels 0;	Gaps 0;

QY		330	TGCGCAATTCCTTTTCTCGCGTGGAAAAATTGCCCTAATGTCTCGAATTCGAAGT	389
Dd		180	TGCCACATCCCATTTACTGTTTTNNNNNNNNNTNNTNNNNNGTNTTAATTTTGTTGT	239
QY		390	TTTTTGCTATGGGTAACTTTTTCCCTATAATTTAATGTTCTTAGGTAAAGATACCTGC	449
Dd		240	TTTTTTTTTTTATTTTAAATATTAATATTTTTTTTTTTTTTTTTTTTTTGATATATTT	299
QY		450	GTCCTACTGTTTTTGTTCAATTTTGTTGCTTTCACC GTTAGTCGCTGATCGGAGTAT	509
Dd		300	TTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGTTTATTT	359
QY		510	TGACGTGAAAACCTCGTTTTTGTTGGTTTTGTTTCATATAAATGAGTATGATCACC	569
Dd		360	TTTTTAATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT	419
QY		570	TTTTTGCTTGGATGTTGTTTTTGAACCTAACGTTGGCGCTGTATATACCTCACG	629
Dd		420	TTTTTGATTTTTTTTTTTTTTTTTTAAATTTTTTTTTTAAATTTTTTTTTTTTTTTTT	479
QY		630	TTCAATGTGAGATTTGAGATTTTGAGTAGTACATGTGGGTTTCTTTGGTGCATAGGTT	689
Dd		480	TTTTTTTTTTTTTAAATTTAATTTTAAATTTAATTTAATTTAATTTAATTTTAAATTT	539
QY		690	GT 691	
Dd		540	AT 541	

```

RESULT 14
US-10-311-455-1931
; Sequence: 1931, Application US/10311455
; Publication No. US20030143606a1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1931
; LENGTH: 14006
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 8289, 8310, 8313
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1931
Query Match      2.0%; Score 66; DB 14; Length 14006;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 14:24:13 ; Search time 182.901 Seconds
(without alignments)
16142.601 Million cell updates/sec

Title: US-10-030-829-1_COPY_1_695
Perfect score: 695
Sequence: 1 gcaacaacaacaataa.....ggcgcgtataggtcgtgaa 695

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002s:*
- 7: geneseqn2003as:*
- 8: geneseqn2003bs:*
- 9: geneseqn2003cs:*
- 10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	695	100.0	3275	4	AAf25373
2	68.8	9.9	6668	6	ABL33697
3	68.6	9.9	556	5	ABV40163
4	68.6	9.9	556	5	ABV40063
5	68.6	9.9	556	5	ABV42105
6	68.6	9.9	556	5	ABV43601
7	66.6	9.6	9539	4	AAa45347
8	66.6	9.6	9539	6	ABK28180
9	66.6	9.5	6644	2	AAK33181
10	66	9.5	7372	2	AAK33182
11	66	9.5	7797	2	AAK33180
12	66	9.5	7996	2	AAK33184
13	66	9.5	14006	6	ABL33958
14	65.8	9.5	556	6	ABQ36997
15	65.8	9.5	556	6	ABQ36996
16	65.6	9.4	113515	6	ABL34174
17	65.4	9.4	6079	6	ABL32421
18	65.2	9.4	887	7	ABT42734
19	64.8	9.3	598	6	ABQ52555
20	64.8	9.3	598	6	ABQ52554
21	64.6	9.3	6668	6	ABL33696
22	64.6	9.3	850	6	ABQ48654
23	64.4	9.3	850	6	ABQ48654

ALIGNMENTS

RESULT 1	AAf25373	AAf25373 standard; DNA; 3275 BP.
AC	AAf25373;	
XX		
DT	15-MAY-2001 (first entry)	
DE	Genomic sequence of the Arabidopsis SG33 gene.	
XX		
KW	SG33 gene; post-transcriptional inactivation; RNA degradation; viral resistance; resistance; fatty acid content; protein content; ss.	
XX		
OS	Arabidopsis thaliana.	
XX		
FH	Key	Location/Qualifiers
FT	exon	696..1658
FT	intron	1659..1731
FT	exon	1732..2023
FT	intron	2024..2134
FT	exon	2135..2379
FT	intron	2380..2481
FT	exon	2482..2648
FT	intron	2649..2738
FT	exon	2739..2949
XX		
PN	MO200105951-A2.	
XX		
PD	25-JAN-2001.	
XX		
PF	13-JUL-2000; 2000MO-FR002052.	
XX		
PR	16-JUL-1999; 99FR-00009417.	
XX		
PA	26-JAN-2000; 2000FR-00001006.	
XX		
XX	(AVET) AVENTIS CROPS SCIENCE SA.	
XX	(INRG) INST NAT RECH AGRONOMIQUE.	

C	24	64.4	9.3	850	6	ABQ48655	Abq48655 Oligonuc1
C	25	64.4	9.3	17934	6	ABL33719	Ab133719 Human imm
C	26	62.8	9.0	529	7	ABX52180	Abx52180 Bovine ES
C	27	62.8	9.0	629	7	ABT21705	Abt21705 Breast ca
C	28	62.8	9.0	700	7	ACD92384	Act92384 Human col
C	29	62.8	9.0	34769	4	AAa46775	Aa46775 Tumour su
C	30	62.6	9.0	578	6	ABQ52047	Abq52047 Oligonuc1
C	31	62.6	9.0	578	6	ABQ52046	Abq52046 Oligonuc1
C	32	62.6	9.0	875	4	AAI95044	Aai95044 Human neu
C	33	62.6	9.0	34769	4	AAa46774	Aa46774 Tumour su
C	34	62.4	9.0	693	6	ABQ32105	Abq32105 Oligonuc1
C	35	62.4	9.0	693	6	ABQ32104	Abq32104 Oligonuc1
C	36	62.4	9.0	1200	6	ABQ39210	Abq39210 Oligonuc1
C	37	62.4	9.0	1200	6	ABQ39211	Abq39211 Oligonuc1
C	38	62.4	9.0	8079	6	ABL92313	Ab192313 Chemical1
C	39	62.4	9.0	23695	6	ABQ66981	Abq66981 Human ang
C	40	62.2	8.9	2357	6	ABN80315	Abn80315 Human che
C	41	62.2	8.9	6334	6	ABL33212	Ab133212 Human imm
C	42	62	8.9	1179	6	ABQ49604	Abq49604 Oligonuc1
C	43	62	8.9	8404	6	ADB54216	Adb54216 Pretrease
C	44	62	8.9	8404	9	ADB54152	Adb54152 Human lym
C	45	62	8.9	8404	9	ADB54152	Adb54152 Human lym


```

QY 336 ATTCTCTTTTCTGCTGGAAAAATGCCCTAATGTTCTCGATTTCGAAGGTTTTGT 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2922 TTTTGTGTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 396 GCTATGGGTACTTTTCCCTATATTTTATAGTCTTAGTAACATACCGCTTA 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2982 TTTTGTGTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 456 CTGTTTTTGTCAATTTGTTGCTTTCACGTTTAGTCGCTAGCGAGTATTACAG 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3042 TTTTGTGTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 516 TGAATAATCCCTGTTTTTGTGTTTGTATATAATAGGATGATACCTTTTGT 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3102 TGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGT 3161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 576 GCTTGAATGTTGTTTGTGAGCCTAAGCGTGTGGCTGTATTAACCTACGTTCA 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3162 TTTTGTGTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 636 TGTGATTTTGAGATTTGGTAGTGAAGTGGGTTCTT 675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3222 TTTTGTGTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3

ABV40163/C

ID ABV40163 standard; cDNA; 556 BP.

AC ABV40163;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 40154.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US005171.

PR 17-FEB-2000; 2000US-018319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 8115; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the CC specification or its complement. (I) is useful for: (a) assessing whether CC a patient is afflicted with prostate cancer; (b) monitoring the CC progression of prostate cancer in a patient; (c) assessing the efficacy CC of a test compound to inhibit prostate cancer; in a patient; (d) assessing CC the efficacy of a therapy for inhibiting prostate cancer in a patient; CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g) CC determining whether prostate cancer has metastasized in a patient; (h) CC assessing the aggressiveness or indolence of prostate cancer in a patient CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker XX

SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

Query Match 9.9%; Score 68.6; DB 5; Length 556;
 Best Local Similarity 48.3%; Pred. No. 8.6e-06;
 Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

QY 265 CTCGCTGTTTACTCTGTAAGTTTCTGCTTAGAGCCTCGCATGCGCTACCGCATGCA 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 486 CMCCTCTCCCTTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 325 TTTCTGCTGATTTCTCTTTTCTGCTGGAATAATGCGCTAATGTTTCGATTG 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 426 TTTTGTGTTTTTTTTTTTATATAAATAAACNCCCTTTTATTTTATTTTATTTT 367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 385 AAGGTTTTTGTGCTAAGGTTACTTTTCCCTAATTTTATAGTCTTAGGTAACGATA 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 366 TTTTGTGTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 445 CCTGCTCTTACTGTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 306 TTTTGTGTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 505 GATTTGACTGTGAAAAATCCCTGCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 246 TTTTGTGTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 565 CTACCTTTTGTGCTTTGATGTTTGTGTTTGTGAGCCTAAGCGTTGCTGTATTAAGT 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186 TTTTGTGTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 625 TCACGTTCAAGTGTGATTTTGAGATTTT 653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 126 TTTTGTGTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4

ABV40063/C

ID ABV40063 standard; cDNA; 556 BP.

AC ABV40063;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 40054.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US005171.

PR 17-FEB-2000; 2000US-018319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1, Page 8100; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC of progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

XX Query Match 9.9%; Score 68.6; DB 5; Length 556;
 XX Best Local Similarity 48.3%; Pred. No. 8.6e-06;
 XX Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

265 CTCGCTGTTACTGCTGTAAGTTTCTGCTTACGCTTCCGATGCGCTTCCGATGCA 324
DB CCCCCCTCCCTTTTATTTTATTTTATTTTCTCCCTCCCTCCCATTAATA 427
QY 325 TTCTGTCGATTCCTTTCTTCTGCTGCAAAATTCCTTAATGTCGATTCG 384
DB 426 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 367
QY 385 AAGCTTTTTCGATGAGGTTTCTTTTCCCTATTTTATGTTCTTGAACGATA 444
DB 366 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 307
QY 445 CTTGCTGCTTACTGTTTGTCTATTTTGTGCTTACCGTTAGTCGATGCGA 504
DB 306 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 247
QY 505 GATTTGATGCTGAAATTCCTGCTTTTGTGTTTGTCTATTAATGAT 564
DB 246 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 187
QY 565 CTACCTTTTGTGCTTGAATGTTGTTTGAAGCTTAAGCTTGTGATTAAGT 624
DB 186 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
QY 625 TCACGTTCAATGTCGATTTTGAATTTT 653
DB 126 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 98
  
```

RESULT 5

ABV42105/C

ID ABV42105 standard; cDNA; 556 BP.

XX AC

XX ABV42105;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 42096.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX OS

XX WO200160860-A2.

XX PN

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183119P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1, Page 8441; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC of progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

XX Query Match 9.9%; Score 68.6; DB 5; Length 556;
 XX Best Local Similarity 48.3%; Pred. No. 8.6e-06;
 XX Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

265 CTCGCTGTTACTGCTGTAAGTTTCTGCTTACGCTTCCGATGCGCTTCCGATGCA 324
DB 486 CCCCCCTCCCTTTTATTTTATTTTATTTTCTCCCTCCCTCCCATTAATA 427
QY 325 TTCTGTCGATTCCTTTCTTCTGCTGCAAAATTCCTTAATGTCGATTCG 384
DB 426 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 367
QY 385 AAGCTTTTTCGATGAGGTTTCTTTTCCCTATTTTATGTTCTTGAACGATA 444
DB 366 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 307
QY 445 CTTGCTGCTTACTGTTTGTCTATTTTGTGCTTACCGTTAGTCGATGCGA 504
DB 306 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 247
QY 505 GATTTGATGCTGAAATTCCTGCTTTTGTGTTTGTCTATTAATGAT 564
DB 246 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 187
QY 565 CTACCTTTTGTGCTTGAATGTTGTTTGAAGCTTAAGCTTGTGATTAAGT 624
DB 186 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
QY 625 TCACGTTCAATGTCGATTTTGAATTTT 653
DB 126 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 98
  
```

RESULT 6

ABV43601/C

ID ABV43601 standard; cDNA; 556 BP.

XX	ABV43601;
XX	16-SEP-2002 (first entry)
DT	
XX	Human prostate expression marker CDNA 43592.
DE	
XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW	pharmacogenomic marker; gene; ss.
XX	
OS	Homo sapiens.
FN	WO200160860-A2.
PD	
XX	23-AUG-2001.
PF	
XX	20-FEB-2001; 2001WO-US005171.
PR	
XX	17-FEB-2000; 2000US-0183319P.
PR	16-MAR-2000; 2000US-0189862P.
PR	25-MAY-2000; 2000US-0207454P.
PR	09-JUN-2000; 2000US-0211314P.
PR	18-JUL-2000; 2000US-0219007P.
PR	13-DEC-2000; 2000US-0255281P.
XX	
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI	Schlegel R, Endege WO, Monahan JE;
XX	
DR	WPI; 2001-662795/76.
XX	
FT	Novel isolated nucleic acid molecule associated with cancerous state of
PT	prostate cells and correlating with presence of prostate cancer, useful
PP	for detecting presence of prostate cancer, stage of prostate cancer.
PS	Claim 1; Page 8681-8682; 11750pp; English.
XX	
CC	The invention relates to an isolated nucleic acid molecule (1) comprising
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC	specification or its complement. (1) is useful for: (a) assessing whether
CC	a patient is afflicted with prostate cancer; (b) monitoring the
CC	progression of prostate cancer in a patient; (c) assessing the efficacy
CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC	the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;
CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC	determining whether prostate cancer has metastasized in a patient; (h)
CC	assessing the aggressiveness or indolence of prostate cancer in a patient
CC	; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX	
SQ	Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;
XX	
Query Match	9.9%; Score 68.6; DB 5; Length 556;
Best Local Similarity	48.3%; Pred. No. 8.6e-06;
Matches 188; Conservative	0; Mismatches 201; Indels 0; Gaps 0
Db	
QY	265 CTCGGTCCTTACTGCTGAAGTTTCTCGCCTTAGAGACCTCCGATGCCCTCACCGCATGCA 324
Db	486 CNCCTCTCCCCCTTTTAATTATTTTTTTTATTTTCTCCCCCCCCTCCCAATAAANA 427
QY	325 TTCTGTGCTGATTTCTCTTTTCTTGCGTGAAAATATGCGCTTAATGTTCTGCATTTTCG 384
Db	426 TTTTATTTTTTTTTTTTTTTTTTAAATAAANCNCCCCCTTTTTTTTTTTTTTTTTT 367
QY	385 AAGGTTTTTGCTAAGGGTACTTTTTTCCCTAATTATTAAGTCTTAGTAAGNANA 444
Db	366 TTTTATTTTTTTTTTTTTTTTATTTTCTTTTTTTTTTTTAAATTTTTTTTTTTTTT 307
QY	445 CCTGGCTTACTGTTTTTGTTTCATTTTTGTGTGCTTTACCGGTTAGCTGATCGGA 504
Db	306 TTTTATTTTTTTTTTTTTTTTTTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTT 247
QY	505 GTAATTGACGTGAAAAATCCTCGTTTGGTTTTGTTTGTTCATATAAATCGAATGAT 564

Dd		246	TTTTTTT	TTTTTTT	TTTTTTT	TTTTTTT	TTTTTTT	TTTTTTT	TTTTTTT	TTCCTTTT	187
Oy		565	CNACCTTTGCGCTTGATGTGGTTTGAAGCCTATGCCTTGATTAACT								624
Dd		186	TT								127
Oy		625	TCACGTTCAATGATGGGATTTGAGATTTT								653
Dd		126	TT								98
<hr/>											
RESULT 7											
AAS45347											
ID	AAS45347	standard; DNA; 9539 BP.									
XX	AC										
XX	AA545347;										
DT	18-DEC-2001	(first entry)									
XX											
DE	Chemically pretreated complementary DNA associated with cell cycle #26.										
XX											
KW	Cell cycle; human; Cpg dinucleotide; cytosine methylation; HIV; aging;										
KW	human immunodeficiency virus; neurodegenerative disorder; solid tumour;										
KW	graft-versus-host disease; glomerular disease; lewy body disease; cancer;										
KW	arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;										
KM	immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; de;										
KX	PCR primer.										
XX											
OS	Homo sapiens.										
PN	WO200168911-A2.										
PD	20-SEP-2001.										
XX											
PF	15-MAR-2001; 2001WO-EP002945.										
PR	15-MAR-2000; 2000DE-01013847.										
PR	06-APR-2000; 2000DE-01019058.										
PR	07-APR-2000; 2000DE-01019173.										
PR	30-JUN-2000; 2000DE-01032529.										
PR	01-SEP-2000; 2000DE-01043826.										
PA	(EPig-) EPIGENOMICS AG.										
PI	Olek A, Piepenbrock C, Berlin K;										
DR	WPI; 2001-602751/68.										
XX											
PT	Designing primers and probes for analyzing diseases associated with										
PT	cytosine methylation state e.g. arthritis, cancer, aging,										
PT	arteriosclerosis comprising fragments of chemically modified genes										
PT	associated with cell cycle.										
PS	Claim 1; SEQ ID NO 52; 28pp; English.										
XX											
XX	Sequences AAS45296-AA545520 represent chemically pretreated genomic DNA										
CC	molecules associated with the cell cycle and specific PCR primers of the										
CC	invention. The sequences are useful for detecting the methylation state										
CC	of all CpG dinucleotides in a sequence and therefore for analysing										
CC	associated diseases. By analysing cytosine methylations in the pretreated										
CC	DNA, genetic and/or epigenetic parameters for the diagnosis and therapy										
CC	of existing diseases or the predisposition to specific diseases can be										
CC	ascertained. The parameters may be compared to another set of genetic										
CC	and/or epigenetic parameters, the differences serving as basis for										
CC	diagnosis and/or prognosis events which are disadvantageous to patients.										
CC	The sequences of the invention are useful for the diagnosis and therapy										
CC	of HIV infection, neurodegenerative disorders, graft-versus-host disease,										
CC	aging, glomerular disease, lewy body disease, arthritis,										
CC	arteriosclerosis, solid tumours and cancers										
QO	Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 U; 0 Other;										

Query Match	9.6%	Score 66.6	DB 4	Length 9539
Best Local Similarity	46.9%	Pred. No. 4.4e-05		
Matches 207; Conservative	0	Mismatches 234	Indels 0	Gaps 0

QY	251	TTTAAACTTACCTCTCCGCTGGTTACACGTAGAGTTTCGCTTAGACCTCGATCG	310
Db	202	TTTTTTTTTTTTGATATTTTTTGGTTATATTTTGTGCTTTTTTTTTTATTTTTTTTC	261
QY	311	CCTCACCGCATCATCTGTGCTGCATTTCTCTTTTTCTTCGCTGAAAAATTGCCCTA	370
Db	262	GTTTTTTTTTTTTTGTGCTGTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTATTT	321
QY	371	TGTTCTCGATTGCAAGGTTTTTGTGCTATGGGTTACTTTTTCCCAATATTATAGTT	430
Db	322	TTATATGTTTTTATTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTT	381
QY	431	CTTAGGTACGADACCGCGCTCACTGTTTTTGTCAATTTGTGTGCTTTCACGGTT	490
Db	382	TTTTTTTTGTTTTTTTTTGTGTTTTTTTTTTTTTTTATTTTTTTTTTGTTTTTTTTTTTT	441
QY	491	AGTCGCTGATCGAGATTTGACTGTGAAAAATCCCTGTTTTTGGTTTTTGTTCA	550
Db	442	ATTTTTTTTTTTTTTATTTTTTTTTTATTTTTTTTTTGTTTTTTTTTTTTTTTTTTATTTT	501
QY	551	TAAATCGATGATCTACTTTTGTGCTTGATGTTGTGTTTTTGAAGCAATGGGTTT	610
Db	502	TTTTTTTTTTTTTGTATTTTTTTTTTTTTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTT	561
QY	611	GGCTGTATTAATCTGACGCTCATGTGAGATTTTGAATTTGGTAGTGAAGTGGGTT	670
Db	562	TTTTTTTTTTGTTTTATTTTTTTTTTTTTTTTTTATTTTATATATTTATTTTTTTTTTTT	621
QY	671	TCTTGGTGCATAGCTGT	691
Db	622	TTATTTATATATATTTTTTTT	642

RESULT 8
ABK28180
ID ABK28180 standard; DNA: 9539 BP.

XX DNA transcription associated complementary genomic DNA #7.
 XX
 XX
 XX DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
 KM PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
 KM single nucleotide polymorphism; adenine deaminase deficiency; cancer;
 KM viral infection; Sezary syndrome; haematological disorder; tuberculosis;
 KM immunological disorder; Werner syndrome; developmental disorder;
 KM porokeratosis; Rieger's syndrome; neurological disorder; erythropoietosis;
 KM neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
 KM myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
 KM angiodysplasia; congenital heart disease; HDR syndrome; gene therapy;
 KM polyglutamine disorder; solid tumour.

OS Unidentified.

PN WO200192565-A2.

06-DEC-2001.

06-APR-2001; 2001WO-EP003973.

06-APR-2000; 2000DE-01019058.

30-JUN-2000; 2000DE-01032529.

	C
F	T
E	G
X	H
	N
	O
	O
	'
	N
	O
	O
	S
	D
	C
	H
	O
	M
	U
	O
	I

33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 105

PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-090046/12.

XX New nucleic acids or oligomers, useful for diagnosing or treating
PT diseases associated with DNA transcription, e.g. immunological disorders,
PT Werner syndrome, psoriasis, myocardial infarction, solid tumors or
PT cancer.

Claim 1; SEQ ID NO 54; 32pp; English.

XX The invention relates to a nucleic acid, which comprises a segment of the
CC chemically pretreated DNA of genes associated with DNA transcription from
CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
CC to the chemically pretreated DNA of genes associated with DNA
CC transcription. The set of oligomer probes are useful for detecting the
CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
CC in a chemically pretreated genomic DNA. The nucleic acids are useful for
CC diagnosing or treating diseases associated with DNA transcription
CC (particularly with the methylation status), e.g. adenoma deamase
CC deficiency, viral infection, retroviral infection, Werner syndrome,
CC haematological disorders, immunological disorders, Menkes syndrome,
CC tuberculous, developmental disorders, psoriasis, Rieger's syndrome,
CC neurodegenerative disorders, neurodegenerative disorders, Waardenburg
CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
CC infarction, hypertension, angiodysplasia, erythropoiesis, congenital heart
CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
CC or cancer. Sequences ABK28137-ABK28472 represent DNA transcription
CC associated genomic DNA molecules of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification but
CC was obtained in electronic format directly from the European Patent
CC Office

Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 U; 0 Other;

Query Match	9.6%	Score 66.6	DB 6	Length 9539
Best Local Similarity	46.9%	Pred. No. 4.4e-05		
Matches 207; Conservative	0	Mismatches 234	Indels 0	Gaps 0

QY	251	TTTAAACTTAGAGTCTCCGTCGTTTACTCTGTGAAGTTTTCGCTTAAGACCTCCGATCG	310
Db	202	TTTTTTTTTTTGTATTTTTCGTTTATATTTTGTGNGTTTTTTTTTTATTTTTTTTTTTC	261
QY	311	CCTCAACCGATGCAATCTGTGTCGATTTCTCTTTTCTTCGCTGGAAAAATTCGCTTA	370
Db	262	GTTTTTTTTTTTTTGTGCGTTTGTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTATTC	321
QY	371	TGTTCTCGATTTTGAAGTTTTTTTGTGCTATAGGTTACTTTTTCCCTATATTTTATAGT	430
Db	322	TATATTTGTTTATTT	381
QY	431	CTTAGGTAAACGATACCTGCGCTTAATCTGTTTTGTTCATTTTGTGTGCTCTCACGTTT	490
Db	382	TTTTTTTTTGTTTTTTTTTCGTTTTTTTTTTTTTTTTTTTATTTTTTTTTTGTTTTTTTTTTTT	441
QY	491	AGTCGCTGAATCGAGATTTTGACTGAAAAAACCCTCGTTTTTGTGTTTGTTCATTA	550
Db	442	ATTTTTTTTTTTTTTTTATTTTTTTTTTTTATTTTTTTTTTTTGTTTTTTTTTTTTTTATTTT	501
QY	551	TAAATCGAATGATCTACCTTTTGTGCTTGATGTTGTTTTGTGAGCGCTATCGCTGT	610
Db	502	TTTTTTTTTTTTTTTGTATTTTTTTTTTTTTTTTTTTTTTGTTTTTTTTTTTTTTTTGTTTTTT	561
QY	611	GGCTTTTAACTTACCTGACGTCATGTGATTTTGAGATTTTGTGATGACTGNGGTT	670
Db	562	TTTTTTTTTTTGTTTTATTTTTTTTTTTTATTTTTTTTAAATAATTTTATTTTTTTTTTTT	621
QY	671	TCTTTGAGGCTATAGGTTGT	691
Db	622	TATTTATATATATTTTTTTT	642

```
RESULT 9
AA33181/c
ID AAX33181 standard; DNA; 6644 BP.
AC AAX33181;
XX
XX
XX 25-JUN-1999 (first entry)
DE Base sequence of the plasmid pRc-ires-bsr.
XX
XX
XX CompoX virus; bsr; viral vector; expression; apoptosis; resistance; crmA;
XX bcl-2; bcl-x1; FLIP; survivin; IAP; ILP; adenovirus; cancer;
XX autoimmune disease; graft rejection reaction; inflammation;
XX inflammatory disease; ss.
XX
XX Synthetic.
XX CompoX virus.
XX WO9913073-A2.
XX
XX 18-MAR-1999.
XX
XX 07-SEP-1998; 98WO-UP004010.
XX
XX 08-SEP-1997; 97JP-00259235.
XX
XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.
XX
XX Hamada H;
XX
XX WPI; 1999-243728/20.
XX
XX New apoptosis-resistant virus-sensitive cell.
XX
XX Example 1; Page 38-41; 51pp; English.
XX
XX The present invention describes an apoptosis-resistant virus-sensitive
XX cell line into which an apoptosis resistance gene has been introduced.
XX The recombinant viruses generated are capable of expressing apoptosis-
XX associated genes. These can then be used in a variety of diseases for
XX which the induction of apoptosis is therapeutic, or where the
XX inhibition of harmful apoptosis is therapeutic. The recombinant viruses
XX are useful as vectors for gene therapy which can be applied to cancer
XX therapy for destroying cancer cells selectively, the treatment of
XX autoimmune diseases and graft rejection reaction, and apoptosis induction
XX therapy for inflammatory cells in inflammatory diseases. Prior arts have
XX encountered the problem where if an adenovirus vector capable of
XX expressing an apoptosis-associated gene is introduced into animal cells,
XX the cells producing the virus will be destroyed because the period of
XX time required to induce cell death by apoptosis is shorter than that
XX required to replicate and produce the virus, resulting in failure to
XX obtain a recombinant virus having the integrated apoptosis-associated
XX gene. In this invention an apoptosis-resistant 293 cell line (having an
XX apoptosis resistant gene introduced) is established and overcomes the
XX problem. The present sequence represents the base sequence of the plasmid
XX pRc-ires-bar, which contains the compoX virus bsr gene, and is used in an
XX example from the present invention
XX
XX Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 U; 0 Other;
SQ
Query Match 9.5%; Score 66; DB 2; Length 6644;
Best Local Similarity 45.4%; Pred. No. 5.6e-05;
Matches 237; Conservative 0; Mismatches 285; Indels 0; Gaps 0;
QY 157 TTTTAGAGCTTAATAAGCTTCTCATTTGTCTCTTCGTCAGTTTATTTCTCTCC 216
DB 11111111111111111111111111111111111111111111111111111
QY 4235 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 4176
DB 11111111111111111111111111111111111111111111111111111
QY 217 GGAGTGCTGACGCTACGCTGCTGCTGCGGCGCTTAACGCTTCGCTGCTTA 276
DB 11111111111111111111111111111111111111111111111111111
QY 4175 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 4116
DB 11111111111111111111111111111111111111111111111111111
QY 277 CTCTGTAAGTTTCTGCTTGAAGCTCGAGTCGCTCAACGCGCATTCCTGCTCGA 336
DB 11111111111111111111111111111111111111111111111111111
```

```
DB 4115 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 4056
QY 337 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 396
DB 4055 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3996
QY 397 CTATGGTACTTTTTCCTTAATTTTATGCTTCTAGGTAACGATACGCTTAC 456
DB 3995 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3936
QY 457 TGTCTTTGTCATTTTGTGCTTTCACCGTTTATGCTGCTAGCGAGTATTGACGT 516
DB 3935 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3876
QY 517 GAAAAATCCTTCGTTTGTGCTTTGCTTTCATATAAATGAGATGATCTACCTTTGTC 576
DB 3875 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3816
QY 577 CTATGATGTTTGTGTTTGTGAGCGCTTACGCTTGTGCTTGTATTAATCTCACGTTCACT 636
DB 3815 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3756
QY 637 GTGATTTTGAGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678
DB 3755 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3714
RESULT 10
AA33182/c
ID AAX33182 standard; DNA; 7372 BP.
XX
XX AAX33182;
XX
XX 25-JUN-1999 (first entry)
XX
XX Base sequence of the plasmid pRc-Bcl-x1-bsr.
XX
XX
XX CompoX virus; bsr; viral vector; expression; apoptosis; resistance; crmA;
XX bcl-2; bcl-x1; FLIP; survivin; IAP; ILP; adenovirus; cancer;
XX autoimmune disease; graft rejection reaction; inflammation;
XX inflammatory disease; ss.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO9913073-A2.
XX
XX 18-MAR-1999.
XX
XX 07-SEP-1998; 98WO-UP004010.
XX
XX 08-SEP-1997; 97JP-00259235.
XX
XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.
XX
XX Hamada H;
XX
XX WPI; 1999-243728/20.
XX
XX New apoptosis-resistant virus-sensitive cell.
XX
XX Example 2; Page 41-45; 51pp; English.
XX
XX The present invention describes an apoptosis-resistant virus-sensitive
XX cell line into which an apoptosis resistance gene has been introduced.
XX The recombinant viruses generated are capable of expressing apoptosis-
XX associated genes. These can then be used in a variety of diseases for
XX which the induction of apoptosis is therapeutic, or where the
XX inhibition of harmful apoptosis is therapeutic. The recombinant viruses
XX are useful as vectors for gene therapy which can be applied to cancer
XX therapy for destroying cancer cells selectively, the treatment of
XX autoimmune diseases and graft rejection reaction, and apoptosis induction
```

CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
CC encountered the problem where if an adenovirus vector capable of
CC expressing an apoptosis-associated gene is introduced into animal cells,
CC the cells producing the virus will be destroyed because the period of
CC time required to induce cell death by apoptosis is shorter than that
CC required to replicate and produce the virus, resulting in failure to
CC obtain a recombinant virus having the integrated apoptosis-associated
CC gene. In this invention an apoptosis-resistant 293 cell line (having an
CC apoptosis resistant gene introduced) is established and overcomes the
CC problem. The present sequence represents the base sequence of the plasmid
CC pR-Bcl-x1-*luc*, which contains the human Bcl-x1 gene, and is used in an
CC example from the present invention

SQ Sequence 7372 BP; 2353 A; 1749 C; 1649 G; 1621 T; 0 U; 0 Other;

Query Match	9.5%;	Score	66;	DB	2;	Length	7372;
Best Local Similarity	45.4%;	Pred.	No.	5.7e-05;			
Matches	237;	Conservative	0;	Mismatches	285;	Indels	0;
						Gaps	0;

OY	157	TTTTAGACTTAATAGCTTCCATAATGTCACCTCCTCGACAGTTAAATTCCTCC	218
Dd	4963	TT	4904
OY	217	GGAGTCCTGACTACTACTCTCACCTCCGGCGCTTAAACTTAGCTCTCCGCCTTA	276
Dd	4903	TT	4844
OY	277	CTGTAGATTCTTGCCCTTAGAGCCTCCGATCCGCTCACGCAATCTGTGCTGA	336
Dd	4843	TT	4784
OY	337	TTTCTCTTTTCTTCGCTGAAAAATGCCCTAATGTCGTGATTCGAAGTTTTGTG	396
Dd	4783	TT	4724
OY	397	CTATGGGTACTCTTTTCCCTAATATTATAGTCTTAGTAAGATACCTGCGCTTAC	456
Dd	4723	TT	4664
OY	457	TGTTTTTGTCATTTGTGTGCTTCAACGCTTAGTGCCTGATCGGAGTATTGACGT	516
Dd	4653	TT	4604
OY	517	GAAAATCCCTGTTTTGTTTTGTTTCATATATAAACGATGATCTACCTTTGNG	576
Dd	4603	TT	4544
OY	577	CTTTGATGTTTTTTGAGCCCTATGCGTGTGGCTGTTATACTTCACGTCATGT	636
Dd	4543	TT	4484
OY	637	GTGGAATTTGAAATTTGGTAGCATGCGGGTTCTTTGT	678
Dd	4483	TTTTTTTTTTTTTTTTTTTGCACCTGCAAGGCGTTATTGGAT	4442

RESULT 11
AA33180/c
ID AA33180 standard; DNA; 7797 BP.

AC AAX33180;

DT 25-JUN-1999 (first entry)

DE Cowpox virus bsr full length gene sequence.

KM Cowpox virus; bsr; viral vector; expression; apoptosis; resistance; crmA;

KW autoimmune disease; graft rejection reaction; inflammation;

XX

XX

PN WO9913073-A2

PD 18-MAR-1999

PF 07-SEP-1998; 98WO-JP004010.

PR 08-SEP-1997; 97JP-00259235.
XX

PA (RPRG-) RPR GENCELL ASIA PACIFIC INC.
XX

PI Hamada H; VY

DR WPI; 1999-243/28/20-
XX

XX Pt New apoptosis-resistant virus-sensitive cells

PS Example 1; Page 34-38; Slipp; English

CC The present invention describes an apoptosis-resistant virus-sensitive
CC cell line into which an apoptosis resistance gene has been introduced.
CC The recombinant viruses generated are capable of expressing apoptosis-
CC associated genes. These can then be used in a variety of diseases for
CC which the induction of apoptosis is by gene transfer, or where the
CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
CC are useful as vectors for gene therapy which can be applied to cancer
CC therapy for destroying cancer cells selectively, the treatment of
CC autoimmune diseases and graft rejection reaction, and apoptosis induction
CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
CC encountered the problem where if an adenovirus vector capable of
CC expressing an apoptosis-associated gene is introduced into animal cells,
CC the cells producing the virus will be destroyed because the period of
CC time required to induce cell death by apoptosis is shorter than that
CC required to replicate and produce the virus, resulting in failure to
CC obtain a recombinant virus having the integrated apoptosis-associated
CC gene. In this invention an apoptosis-resistant 293 cell line having an
CC apoptosis resistant gene introduced) is established and overcomes the
CC problem. The present sequence represents the complex virus per gene which
CC is used in an example from the present invention

SQ Sequence 7797 BP; 2542 A; 1760 C; 1656 G; 1839 T; 0 U; 0 Other;

Query Match	9.5%;	Score 66;	DB 2;	Length 7797;
Best Local Similarity	45.4%;	Pred. No. 5.8e-05;		
Matches 237;	Conservative	0;	Mismatches 285;	Indels 0;
				Gaps 0;

[illegible]

QY 577 CTTTGATGTTGTTTGGAGCCGATGCGTGGCTGTTAACTCAAGTCATGT 636
 Db 4968 TTTTGT 4908
 QY 637 GTGATTTTGGATTTTGGTAGTCAGTCGTGGTTCTTTGGT 678
 Db 4908 TTTTGT 4867
 RESULT 12
 ID AAX33184/c
 ID AAX33184 standard; DNA; 7996 BP.
 XX AAX33184;
 AC
 XX
 XX 25-JUN-1999 (first entry)
 DE
 XX Base sequence of the plasmid pRX-Bcl 2-i-hcd 25.
 XX
 XX
 KW Compox virus; bsr; viral vector; expression; apoptosis; resistance; crmA;
 KW bcl-2; bcl-xl; FltP; survivin; IAP; ILP; adenovirus; cancer;
 KW autoimmune disease; graft rejection reaction; inflammation;
 KW inflammatory disease; ss.
 XX
 XX Synthetic.
 OS
 OS Homo sapiens.
 XX
 PN W09913073-A2.
 PN
 PD 18-MAR-1999.
 XX
 PF 07-SEP-1998; 98WO-JP004010.
 XX
 PF 08-SEP-1997; 97JP-00259235.
 XX
 XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.
 PA
 PI Hamada H;
 XX
 XX WPI; 1999-243728/20.
 DR
 XX
 XX
 XX
 PS New apoptosis-resistant virus-sensitive cell.
 PS Example 3; Page 46-49; 51pp; English.
 XX
 XX The present invention describes an apoptosis-resistant virus-sensitive
 CC cell line into which an apoptosis resistance gene has been introduced.
 CC The recombinant viruses generated are capable of expressing apoptosis-
 CC associated genes. These can then be used in a variety of diseases for
 CC which the induction of apoptosis by gene transfer, or where the
 CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
 CC are useful as vectors for gene therapy which can be applied to cancer
 CC chemotherapy for destroying cancer cells selectively, the treatment of
 CC autoimmune diseases and graft rejection reaction, and apoptosis induction
 CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
 CC encountered the problem where if an adenovirus vector capable of
 CC expressing an apoptosis-associated gene is introduced into animal cells,
 CC the cells producing the virus will be destroyed because the period of
 CC time required to induce cell death by apoptosis is shorter than that
 CC required to replicate and produce the virus, resulting in failure to
 CC obtain a recombinant virus having the integrated apoptosis-associated
 CC gene. In this invention an apoptosis-resistant 293 cell line (having an
 CC apoptosis resistant gene introduced) is established and overcomes the
 CC problem. The present sequence represents the bse sequence of the plasmid
 CC pRX-Bcl 2-i-hcd 25, which contains the human Bcl-2 gene, and is used in
 CC an example from the present invention
 XX
 XX Sequence 7996 BP; 2463 A; 2015 C; 1829 G; 1689 T; 0 U; 0 Other;

QY	157	TTTTAGAGCTAAATAAGCTTCCCATTTGCTCTCTCCGAGTTAATTTCTCTCC	216
Db	5587	TT	5528
QY	217	GGAGTCCGCACTACACTCTCACTCTCCGCGCTTTAACTTACGTTCTCCGTCTTA	276
Db	5527	TT	5468
QY	277	CTCTGAGATTTTCTGCGCTTAGAGCCCTCGATCGCCTACCGCARGCATCTGTGCTGA	336
Db	5467	TT	5408
QY	337	TTTCTCTTTTCTCTCGCTGAGAAAATGCCCTAATGCTTCGATTTGGAAGTTTTGTG	396
Db	5407	TT	5348
QY	397	CTATGGGTACTTTTTTCCCTAATTTTATGTTCTTAGGAAAGATACCGCTCTTAC	456
Db	5347	TT	5288
QY	457	TGTTTTTGTTCAATTTGTGTGCTTTCACCGTTTAGTGTGATCGAGATTTGACTGT	516
Db	5287	TT	5228
QY	517	GAAAAACCTCGCTTTTGTGTTTGTCTTCATATAATGGATGATCACTTTGTG	576
Db	5227	TT	5168
QY	577	CTTGTAGTTTGTGTTTTTGAGACCTAGTGTGTGCTGTATATACCTGACTCANGT	636
Db	5167	TT	5108
QY	637	GTCGATTTTGAGATTTTGTGTAGTGCATGTGGGTTTCTTTGGT	678
Db	5107	TTTTTTTTTTTTTTTTTTGCAACGTGACAGAGGGTTTATTTGAT	5066
RESULT 13			
ABL33958			
XX	ABL33958	standard; DNA; 14006 BP.	
AC	ABL33958;		
XX			
DT	26-MAR-2002	(first entry)	
XX			
DE	Human immune system associated gene SEQ ID NO: 1931.		
XX			
KM	Human, immune system disease; cytosine methylation; antiasthmatic;		
KM	antiarteriosclerotic; antianaemic; cytosolic; noctropic;		
KM	neutroprotective; anti-HIV; anticonvulsant; ophthalmological;		
KM	antipneumatic; antiarthritic; antidiabetic; antipsoriatic;		
KM	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;		
KM	acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;		
KM	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;		
XX	ds.		
XX			
OS	Homo sapiens.		
XX			
PN	MO200200928-A2.		
XX			
PD	03-JAN-2002.		
XX			
PF	02-JUL-2001; 2001MO-EP007537.		
XX			
PR	30-JUN-2000; 2000DE-01032529.		
XX			
PR	01-SEP-2000; 2000DE-01043826.		
XX			
PA	(EPIG-) EPIGENOMICS AG.		
XX			
P1	Olek A, Piepenbrock C, Berlin K;		
XX			
DR	WPI; 2002-130909/17.		
XX			

RESULT 15

AB036996 standard; DNA; 556 BP.

AB036996;

12-JUL-2002 (first entry)

Oligonucleotide for detecting cytosine methylation SEQ ID NO 23587.

Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;

drug; side effect; cancer; central nervous system; cardiovascular;

gastrointestinal; respiratory system; single nucleotide polymorphism;

SNP; cell differentiation; ds

Homo sapiens.

WO200218632-A2.

01-SEP-2001; 2001WO-EP010074.

01-SEP-2000; 2000DE-01043826.

05-SEP-2000; 2000DE-01044543.

(EPIC-) EPIDENOMICS AG.

Olek A, Piepenbrock C, Berlin K, Gueitig D;

WPI; 2002-371829/40.

Determining the degree of cytosine methylation in genomic DNA, useful for

diagnosis and prognosis, comprises selective hybridization of amplicons

from chemically treated DNA.

Claim 12; 56pp + Sequence Listing; 56pp; German.

This invention describes a novel method for determining the degree of

methylation of a particular cytosine in a motif 5'-CpG-3', present in a

genomic sample of DNA. The sample is treated chemically to convert

cytosine (C) but not methylated C, to uracil, then part of the genomic

DNA that contains the target C is amplified to form a labeled amplicon.

The amplicon is hybridised to two classes, each with at least one member,

of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the

degree of hybridisation to both classes is determined from the label on

the amplicon. From the ratio of labels hybridised to the two classes of

oligomers, the degree of methylation is calculated. The method is used:

(i) for diagnosis and/or prognosis of side effects of therapeutic drugs

and of a wide range of diseases, e.g. cancer, disorders of the central

nervous, cardiovascular, gastrointestinal and respiratory systems etc.,

particularly by detecting mutations or single-nucleotide polymorphisms

(SNP/s); and (ii) for differentiation of cell or tissue types and for

investigating cell differentiation. The method allows the methylation

status of many C residues to be determined simultaneously. AB013410-

AB054121 represent genomic DNA sequences used to illustrate the method

for determining the degree of cytosine methylation described in the

disclosure of the invention

Sequence 556 BP; 12 A; 13 C; 14 G; 505 T; 0 U; 12 Other;

Query Match 9.5%; Score 65.8; DB 6; Length 556;

Best Local Similarity 49.3%; Pred. No. 3.6e-05;

Matches 172; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

325 TTCTGAGCTGATTTCTCTTTTTCGCGTGAATAATGCCCTAAGTTCGATTCG 384

96 TTTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTATTTTTTTTTTTT 155

385 AAGGTTTTTGCTAGGCTTACTTTTTTCCCTATATTTATATTTCTTAGGTAACGATA 444

Db 156 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTACGTTTTTTTTTCGTTTTTTTTTTTTT 215

Qy 445 CCGGCGCTTACGTTTTTTGTTCAATTGTTGNGCGTTTCACCGTTTAGTCGTCGATTCGA 504

Db 216 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTTTT 275

Qy 505 GTATTTCACCTGAAAAATCCCTGTTTTTTGTTTTGTTTCATATAAATCGGATTCGAT 564

Db 276 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTTTT 335

Qy 565 CTACCTTTTTCGCTTTCGATGTTGTTTTTTGAGCCATAGCGTTTCGCTGTGTAATCT 624

Db 336 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTTTT 395

Qy 625 TCACGTTATGTTGATTTTGAATTTTGTAGTACTGTGGTTCT 673

Db 396 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTCGTTTTTTT 444

Search completed: April 8, 2004, 09:18:38

Job time : 184.901 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 19:26:20 ; Search time 1216.13 Seconds
(without alignments)
17065.771 Million cell updates/sec

Title: US-10-030-829-1_COPY_1_695
Perfect score: 695
Sequence: 1 gcaacaacaacaataa.....ggctgctatagctgtaaa 695

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pln:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_phg:*
27: em_gse_vrt1:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	364.6	52.5	395	29	BX547725 Arabidops
2	277.4	39.9	303	29	BX547724 Arabidops
3	157.8	22.7	209	28	BH854770 SALK_0886
4	108.8	15.7	144	28	BH847287 SALK_0508

5	103.4	14.9	232	28	AQ964581
6	92	13.2	706	14	CF872119
7	88.4	12.7	922	29	CNS00731W
8	88	12.7	997	29	CNS005TE
9	87.2	12.5	712	13	BX416727
10	84.8	12.2	613	28	AQ964580
11	84	12.1	647	9	AU236368
12	82.8	11.9	172	28	BZ558518
13	81.2	11.7	1068	29	CNS008TV
14	81	11.7	81	28	BH791768
15	80.4	11.6	1201	9	AL514657
16	79.8	11.5	1295	29	CG757539
17	79.6	11.5	922	28	AZ548363
18	77.8	11.2	1180	13	BX436369
19	76.2	11.0	1201	13	BX345077
20	76.2	11.0	1321	29	CG746535
21	76	10.9	1101	29	CNS00LT2
22	75.8	10.9	1198	28	B08337
23	75.8	10.9	1528	29	CG753854
24	75	10.8	1224	28	CG232115
25	75	10.8	1274	29	CG745106
26	74.8	10.8	896	29	CNS008P8
27	74.6	10.7	1201	13	BX332305
28	74.6	10.7	1433	29	CG745119
29	74.4	10.7	889	29	CNS006W4
30	74.4	10.7	961	29	AG072942
31	74.4	10.7	1101	29	CNS008WD
32	74.2	10.7	894	9	AL569690
33	74.2	10.7	949	12	BM416116
34	74.2	10.7	962	14	CD050907
35	74.2	10.7	1101	29	CNS003B0
36	74.2	10.7	1269	29	CG749445
37	74	10.6	775	29	AG061601
38	74	10.6	870	28	AQ866797
39	73.8	10.6	846	13	B0745382
40	73.8	10.6	1286	28	CC240677
41	73.6	10.6	879	13	BX456766
42	73.6	10.6	893	28	BZ690598
43	73.4	10.6	839	29	CNS0122R
44	73.2	10.5	731	9	AI557572
45	73.2	10.5	817	29	CNS009FW

ALIGNMENTS

RESULT 1	BX547725	395 bp	DNA	linear	GSS 02-JUL-2003
LOCUS	Arabidopsis thaliana T-DNA flanking sequence GK-547E01-020631.				
DEFINITION	Genomic survey sequence.				
ACCESSION	BX547725				
VERSION	BX547725.1	GI:32440534			
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H. and Weishaar,B.				
AUTHORS	A pipeline for automated high-throughput generation of Fsts (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines				
TITLE	Unpublished				
JOURNAL	Rosso,M., Strizhov,N., Li,Y., Reis,B., Dekker,K. and Weishaar,B.				
REFERENCE	A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)				
AUTHORS	for flanking sequence tag based reverse genetics				
REFERENCE	3 (bases 1 to 395)				
AUTHORS	Strizhov,N., Li,Y., Rosso,M. and Weishaar,B.				

TITLE	JOURNAL	COMMENT
Direct Submission	Submitted (01-JUL-2003) Weishaar B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany	This sequence is recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by clone MW1. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: http://www.mpiz-koeln.mpg.de/GABI-Kat/ .

Query Match	Similarity	52.5%	Score 364.6;	DB 29;	Length 395;
Best Local	Similarity	95.2%;	Pred. No. 2,1e-50;		
Matches	Conservative	0;	Mismatches 19;	Indels 0;	Gaps 0;
QY	114	ACTGATGAAAAATTTGGATCGAATCGAAAAACGAGCCGTTTATGAGCTTAATAG	173		
Db	1	ACGATGAAAAATTTGAGATCGAATCGAAAAACGAGCCGTTTAAAGCTTAATAG	60		
QY	174	CTTCCGATTTGCTCTTCTGTCAGTTATTTTCTCTCGAGTCTGACTACACTA	233		
Db	61	CTTCCGATTTGCTCTTCTGTCAGTTATTTTCTCTCGAGTCTGACTACACTA	120		
QY	234	CTCTCACTCCGCGCGCTTAAACTTAAGTTCGTCGTCGTTTACTGTAAATTTCTCG	293		
Db	121	CTCTCACTCCGCGCGCTTAAACTTAAGTTCGTCGTCGTTTACTGTAAATTTCTCG	180		
QY	294	CTTAAGACCTCCGATCGCTCACCGCATGAGTCTGTGTCGATTTCTTTTCTTCGC	353		
Db	181	CTTAAGACCTCCGATCGCTCACCGCATGAGTCTGTGTCGATTTCTTTTCTTCGC	240		
QY	354	TGGAAAAATGACCCTAATGTTCTCGATTCGAAAGTTTTGTGCTATGAGTTACTTTTT	413		
Db	241	TGGAAAAATGACCCTAATGTTCTCGATTCGAAAGTTTTGTGCTATGAGTTACTTTTT	300		
QY	414	CCCTAATATTTAATGTTCTTAAGTAAAGTAACCTGCGCTTTACGTTTTGTGCATTTG	473		
Db	301	CCCTAATATTTAATGTTCTTAAGTAAAGTAACCTGCGCTTTAATGTTTTGTGCATTTG	360		
QY	474	TTTGCTTTACCGTTTATGTCGTCGATCGAGTAT	508		
Db	361	TTTGCTTTACCGCTATTTGTGCTGATCAGAGAT	395		

RESULT 2	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
BX547724	BX547724	303 bp DNA	linear	GSS-02-JUL-2007			
	Arabidopsis thaliana T-DNA flanking sequence GK-547E01-020586,						
	genomic survey sequence.						
	BX547724						
	BX547724.1	GI:32440533					
	GSS.						
	Arabidopsis thaliana (thale cress)						
	Arabidopsis thaliana						
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						

REFERENCE AUTHORS	TITLE
1 Stetznov, N., Li, Y., Rosso, M., Viehoveer, P., Dekker, K., Saedler, H. and Weisshaar, B.	A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA

2
Rosso, M., Srichoy, N., Li, Y., Reis, B., Dekker, K. and Weishaar, B
A new *Arbidopsis thaliana* T-DNA mutagenised population (GABI-Kat)
for flanking sequence tag based reverse genetics

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
3 (bases 1 to 303)
Rosso, M., Li, Y., Strizhov, N. and Weishaar, B.
Direct Submission
Submitted (01-JUN-2003) Weishaar B., Max-Planck-Institut fuer
Chemie, Carl-Neuberg-Strasse 1, D-50939, Koeln, 50829, Germany

FEATURES
SOURCE

ORIGIN				
Query Match	39.9%;	Score 277.4;	DB 29;	Length 303;
Best Local Similarity	94.7%;			
Best Match	287;	Pred. No. 5e-36;		
Matches	287;	0;	Mismatches 16;	Indels 0;
				Gaps 0;

QY	415	CCT	417
Db	301	CCT	303

RESULT 3
BH854770
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES

BH854770 209 bp DNA linear GSS 08-JUL-2002
SALK_088648.19.55.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_088648.19.55.x, genomic
survey sequence.
BH854770
BH854770.1 GI:21704360
GSS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 209)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadriab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shim,P., Zimmerman,J. and Ecker,J.R.
A Sequence-indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
Location/Qualifiers
1..209
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_088648.19.55.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 22.7%; Score 157.8; DB 28; Length 209;
Best Local Similarity 84.7%; Pred. No. 2.6e-16;
Matches 177; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 322 GCATTCTGCTCGATTTCTTTTCTTCTGCGTGAATAATGCGCTTAATGTTCTGATT 381
|||||
DB 1 GCATTCTGCTCGATTTCTTTTCTTCTGCGTGAATAATGCGCTTAATGTTCTGATT 60
|||||
QY 382 TCGAAGGTTTGTGCTATGCGTTACTTTTCCATATTTATTTAGTTCTTAAGTAAG 441
|||||
DB 61 TCGAAGGTTTGTGCTACGCGTTACTTTTCCATATTTCTTAATGTTAGTGGGCGG 120
|||||
QY 442 ATACCTGCGTCTTACTGTTTGTTCATTGTTGTGCTTTCACCGTTTAGTCGTGATC 501
|||||
DB 121 ATGCCGCGCTTTAGTGGCTATGACAGTTGGGGGCGCTTTCCCGGTTAGCGCGTATC 180
|||||
QY 502 GGAATATTGACTGTGAAAAATCTTCTGCT 530
|||||
DB 181 GGAAGATTGACTGGGTAATAATCTTCTGCT 209
|||||

RESULT 4
BH847287
LOCUS
DEFINITION

BH847287 144 bp DNA linear GSS 13-JUN-2002
SALK_050870.18.25.x Arabidopsis thaliana TDNA insertion lines

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES

BH847287 232 bp DNA linear GSS 28-JUN-2000
SALK_050870.18.25.x Arabidopsis thaliana genomic clone LERGX20, genomic
survey sequence.
BH847287
BH847287.1 GI:21418158
GSS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 144)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadriab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shim,P., Zimmerman,J. and Ecker,J.R.
A Sequence-indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
Location/Qualifiers
1..144
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_050870.18.25.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 15.7%; Score 108.8; DB 28; Length 144;
Best Local Similarity 84.7%; Pred. No. 3.6e-08;
Matches 122; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 372 GTTCTCGATTGGAAGTTTGTGCTATGAGTTACTTTTCCATATTTATTAAGTTC 431
|||||
DB 1 GTTCTCGATTGGAAGTTTGTGCTATGAGCGCTACTTTTCCATATTTATTAAGTTC 60
|||||
QY 432 TTAGTAAGATACCTGCGTCTTACTGTTTGTTCATTGTTGTGCTTTCACCGTTTA 491
|||||
DB 61 TTAGTAAGATACCTGCGTCTTACTGTTTGTGCGCTTGTGCGATTGGCGCGCTGA 120
|||||
QY 492 GTGCGTATCGAGTATTGACTG 515
|||||
DB 121 GTGCGTATCGAGGATATTGACTG 144
|||||

RESULT 5
A0964581
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

A0964581 232 bp DNA linear GSS 28-JUN-2000
LERGX20TR LERX Arabidopsis thaliana genomic clone LERGX20, genomic
survey sequence.
A0964581
A0964581.1 GI:6792282
GSS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE	AUTHORS
1 (bases 1 to 232)	Pai, G., Barnstead, M., Bowman, C., Utterbach, T., Buell, C. R., Lin, X., and Fraser, C. M.

TITLE
Genomic survey sequencing of Landsberg erecta ecotype of
Arabidopsis thaliana and identification of sequence-based

JOURNAL COMMENT
Unpublished (2000)
Contact: Xiaoying Lin
mailto:linxiaoying@econ.au.dk
Research Institute for Economic Research

For additional information, see <http://www.tigr.org/tcb/at/at.html>.
Seq primer: TR
Class: shotgun.

ORIGIN

Query Match	14.9%	Score 103.4	DB 28	Length 232
Best Local Similarity	94.7%	Pred. No. 2e-07		
Matches 107, Conservative	0	Mismatches 6	Indels 0	Gaps 0

Oy 583 TGTTCGTTTTTTACGCCCTAATGCCTGTTGGCTTGTTATACTTCAAGTTCATGTGTGGAT 64
| | | | |
Db 16 TGCTGGAAGTTGAGCCTATGCGTTGGCTTGTTATAACTTCAAGTTCATGTGTGGAT 75

```
QY      643 TTTGAGATTTTGGTAGTGACTGTGGGTTCTTTGGTGGCATAAGTTGTAAAA 695
        |||||
DB      76   TTTGAGATTTTGGTAGTGACTGTGGGTTCTTTGGTGGCATAAGTTGTAAAA 128
```

RESULT 6			
CF872119			
LOCUS	706 bp	mRNA	linear
DEFINITION	trf:c02xrp16.b1 T. reesei mycelial culture, Version 6 October 2003		
	Hypocrea jecorina cDNA clone trfc02xrp16,	mRNA sequence.	

SOURCE	ORGANISM
Hypocrea jecorina (anamorph: <i>Trichoderma reesei</i>)	Hypocrea jecorina
Hypocrea jecorina	Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes	Hypocrea jecorina
Hypocrymmycetidae; Hypocreales; Hypocreaceae; Hypocrea	Hypocrea jecorina

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 706)	Diener, S.E., Dankmeyer, L., Dunn-Coleman, N., Houfek, T.D., Mitchell, T.K., van Solingen, P., Teunissen, P.J.M., Ward, M. and Dean, R.A.	Analysis of the protein processing and secretion pathways in a <i>Trichoderma reesei</i> EST dataset	Unpublished (2003)

COMMENT

FEATURES

```

/db xref="taxon:51453"
/clone="rric029xp16"
/dev stage="mycelia"
/clone_1lb="T.reesei mycelial culture, Version 6 October
2003"
_note="Vector: pRR31; Site 1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."

```

ORIGIN

Query Match	13.2%	Score 92;	DB 14;	length 706;
Best Local Similarity	45.9%	Pred. No. 6.9e-06;		
Matches 308; Conservative	0;	Mismatches 363;	Indels 0;	Gaps 0

Qy
5 AACCAACAAAATTAGCAGTATGTTCTGACATAAATTATAGTGGAACATTA 64

Ddb
4 AA 63

Oy 65 GTTAAAGCCGAAAAAGAAAAAGGTACAAAATGAAAAAATACTCAACTGAATGAAA 124
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 AA 123

```

QY      125 ATTTGGAGTCCAGAAATCGGAAAAAGAGCGCGCTTTAGAGCTTAATAAGCTTCCTCATTT 184
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      124 AAAAGGAGGCGCGCGGTTTCNTTTTTTTTTTTTTTTCTTTTTCCTTTTTCCTCTT 183

```

Qy 185 GTCCTTCTTCGACGATTATTTCTTCCGCCGAGTCCTGCATCTACTCTCACTCTC 244
| | | | | | | | | | | | | | | | | |
Db 184 TTTTTTTTTTTTTTTTTTTTTTTTTTTCTTTTCTTTTTTTTTTTTTTTTTTTT 243

Qy . 245 CGGCGCTTAAACCTTAGCTTCCTCCGTCTTACTGTAGTTTTCGCCTTAGAGCTC 304
||| ||| ||| ||| ||| ||| ||| |||
Db 244 TTTTCCTTTCCTTTTTTTTTTCTTTTTCCTTTCCTTTCTTTTCTTTTCTTTTCTC 303

```

QY      305 CGATCGCCTGACCGCATGATTCGTGTCTGATTCCTTTTCTTCGGCTGGAATAATG 364
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      304 TTTTTCCTTTTCCTTTTCCTTTTCCTTTTCCTTTTCCTTTTCCTTTTCCTTTTC 363

```

QY	365	CCCTAATGTTCTCGATTTCGAAGGTTTTTGTCATAGGGTACCTTTTTCCCTATATTTT	424
Db	364	CTTTTTCTTTTTTTTTCTTTCTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	423

Dy 425 ATAGTCTTAGTAACGATACCTGGCTTTACTGTTCATTTCATTTTGTTGCCTTCA 484
| | | | | | | | | | | | | | | |
Db 424 TT 483

QY 485 CCGTTTAGTCGCTGATCGAGTATTGACTGTGAATAATCCTTCGTTTTTTGGTTTTGT 544
||| | | | | | | | | | | | |
Db 484 TTTTITTTTTCTTTTTTTTTTTTTTTTTCTTTCTTTCTTTCTTTCTTTCTTTT 543

Dy
Db

545 TTCAATAAATCGATTGACTCTTTGTGGTTCAGTTTTTGACCCCTTAGC 604
||| || | || | || | || | || |
544 TTTTITTTTCCTTTNNNTTTTITTTTCCTTCTTTTCTTTTCTTTTCTTTTCTTTT 603

DY 605 GTTGTGGCTTGTTAATCACTTCACGTCATGTGTGATTGGATTTGGTAGACTG 664
||| ||| ||| ||| ||| ||| ||| |||
DB 604 TTTTTTTCTTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTCCCTTTCTTTT 663

QY	665 TGGGTTTCTTT	675
Db	664 CTTTTTTTTTT	674

LOCUS	922 bp	DNA	linear	GSS 03-JUN-1995
CNS0073W				
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #			
	BAC11409 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			

ACCESSION	AL066784
VERSION	AL066784.1
KEYWORDS	GI:4945247
SOURCE	GSS. <i>Drosophila melanogaster</i> (Fruit fly)

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 922)

REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..922
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14D09"
/clone_1ib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 12.7%; Score 88.4; DB 29; Length 922;
Best Local Similarity 29.4%; Pred. No. 2.3e-05;
Matches 200; Conservative 165; Mismatches 309; Indels 6; Gaps 2;

QY 2 ACAACAACAAATTAAGCAGTCATGCTGAGCAATTAATTAAGTGGAAACAT 61
DB 239 AAAAAACWCTAACTAATAATTAACAAATAMCAACAAATTAATTAACAAAAACAA 298
QY 62 TAGTTAAGCGAAGAAAGAAAGAAAGTACAAATGAAGAAACAAATGAAGTGAAG 121
DB 299 CAAAAAACCAAAAAAAN 358
QY 122 AAAATTGAGTCAGAAATCGAAAAACGAGCGCTTTAGAGCTTAATTAAGCTTCTCA 181
DB 359 AAAA-----AACAAAAAATAAATAAACCCCAAAAAAACACATCCTCAATCTTT 413
QY 182 TTGTCTCTTCTGTCAGTTAATTTCTCTCCGAGTCGACTCACTCACTCACT 241
DB 414 CTTCGCTTCCCTCTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 473
QY 242 CTCGGCGGCTTAACCTAAGTCTCCGCTGCTTAAGTCTGTAAG-TTTCTGCTAGAG 300
DB 474 CTGGGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTG 533
QY 301 CCTCCAGTCGCTCACCGCAGTCTGTCGATTTCTCTTTTCTCTGCTGGAATA 360
DB 534 GGTGKGKGTGTTTSTGKGKGTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTG 593
QY 361 ATTGCCCTAATGTTCTGATTCGAAAGTTTGTGCTAGTGGTTACTTTTCCCTA 420
DB 594 TGGGKGKGTGTTTGTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTG 653
QY 421 TTTTATAGTTCTTAAGTAAGTACCTGCTTAAGTCTTATGTTGCTATTTGTTGCT 480
DB 654 TTGGGKKTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKTGKT 713
QY 481 TTGACGCTTATGCTGATCGAGATTTGACGTGAAAAATCTTCTGTTTGTGTTT 540
DB 162 GAGCTTAATAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 221

DB 714 TTKTTKTKKKKTK 773
QY 541 TTGTTTCATATAATGAGTATGATACCTTTTGCTTGAATGTTGTTTTCAGCCT 600
DB 774 KGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKG 833
QY 601 ATGCGTTGTTGCTGTTTAACTTCACGTTGATGATGATGATGATGATGATGATG 660
DB 834 GTKTKGTGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKG 893
QY 661 ACTGTGGGTTCTTTGCTG 680
DB 894 KGKKGKKGKKGKKGKKGKKG 913

RESULT 8
CNS005TE 997 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR12K22 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL060767.1 GI:4943573
VERSION AL060767.1
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 997)

REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..997
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12K22"
/clone_1ib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 12.7%; Score 88; DB 29; Length 997;
Best Local Similarity 20.8%; Pred. No. 2.5e-05;
Matches 115; Conservative 160; Mismatches 278; Indels 0; Gaps 0;

QY 42 AAATTAATGAGGGAACAATTAAGTAAAGCAAAAAAGAAAAAGGTACAAAAATGA 101
DB 335 ATATNAANNNAN 394
QY 102 AAACAAATCAACGAATGAATTTGAGATCGGAATCGAAAAACGAGCGCTTTTA 161
DB 395 ANAAAAANANNTATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 454
QY 162 GAGCTTAATAGCTTCT 221

[illegible]

RESULT	9
LOCUS	BX416727
DEFINITION	BX416727 Homo sapiens NEUROBLASTOMA Homo sapiens CDNA clone
ACCESSION	BX416727
VERSION	BX416727.1 GI:30765629
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Hominiidae; Homo. 1. (bases 1 to 712) Li, W.B., Gruber, C., Jesssee, J. and Polayes, D. Full-length cDNA libraries and normalization Unpublished (2001)
JOURNAL	Contact: Genoscope
COMMENT	Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr library was constructed by Life Technologies, a division of Invitrogen. Contact : Peng Liang Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradey Avenue Genoscope sequence ID : CSDDA011BE07QPl.
FEATURES	
source	1..712 Location/Qualifiers

```

/db_xref="taxon:9606"
/clone="CSODNA011Y114"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

Matches	54;	Conservative	185;	Mismatches	129;	Indels	0;	Gaps	0;
QY	324	ATTCTGCGTCGATTTCTTTTCTCTGCGTGAAAAATGCCCTAATGTTCTGATTC	383						
	:: ::								
Db	330	MMMT	389						
	:: ::								
QY	384	GAAGTTTGTGCTAGGTTACTTTTCCATATTTATAGTTCTAGTAAGAT	443						
		:: ::							
Db	390	TT	449						
		:: ::							
QY	444	ACGCGGCTTACGTTTGTGTCATTTGTGTCGTTGACGTTGCGTGATCG	503						
	:: :: :: :: :: :: :: :: :: :: ::	:: :: :: :: :: :: :: :: :: :: ::							
Db	450	KTTTKTKTKKKKKTKTKTKTKGKKKKKKKKKKKKKKKKKKKKKKKKKKKK	509						
		:: ::							
QY	504	AGTATTGACGCGTAAAAATCTTCGTTTTTGTGTTTTGTTTCATATAATCGATTGA	563						
	:: :: :: :: :: :: :: :: :: :: :: ::	:: :: :: :: :: :: :: :: :: :: ::							
Db	510	KKK	569						
		:: ::							
QY	564	TCTACCTTTTGTGCTTGATGTTGTTTTTTTGAACCTATGCGTTGTGATTAAC	623						
	:: :: :: :: :: :: :: :: :: :: :: ::	:: :: :: :: :: :: :: :: :: :: ::							
Db	570	KKK	629						
		:: ::							
QY	624	TTACGTTTCATGTGATTTTGTGATTTTGTAGATGACTGTGGATTTCCTTGTGAGTA	683						
	:: :: :: :: :: :: :: :: :: :: :: ::	:: :: :: :: :: :: :: :: :: :: ::							
Db	630	KKK	689						
		:: ::							
QY	684	TAGGTTGT 691							
	:: :: :: :: :: :: :: :: :: :: :: ::	:: :: :: :: :: :: :: :: :: :: ::							
Db	690	KKKKKKK 697							
		:: ::							

```

RESULT 10
A0964580/c
LOCUS
DEFINITION A0964580 613 bp DNA linear GSS 28-JAN-2000
LERRGX20TF LERG Arabidopsis thaliana genomic clone LERRGX20, genomic
survey sequence.
ACCESSION
A0964580
VERSION
A0964580.1 GI:6792281
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 613)
Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Uteerbach,T.,
Reidblyum,T., Liang,F., Creasy,T. and Frazer,C.M.
Genomic survey sequencing of Landsberg erecta ecotype of
Arabidopsis thaliana and identification of sequence-based
polymorphisms
unpublished (2000)
JOURNAL
COMMENT
Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: at.tigr.org
For additional information, see http://www.tigr.org/cdb/at.html
Seq primer: TF
Class: shotgun.
FEATURES
source
1..613
location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="LANDSBERG ERECTA"
/db_xref="taxon:3702"
/clone="LERGX20"
/clone_lib="LERG"
/notes="Organ: Leaf; Vector: pUC19K; Total genomic DNA was
sheared to 0.4-0.7 kbp before ligation."

```

ORIGIN	Query Match	12.2%	Score 84.8	DB 28	Length 613
--------	-------------	-------	------------	-------	------------

Best Local Similarity 97.7%; Pred. No. 0.00012;
Matches 86; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 608 GTTGGCTGTTTAACTTCACGTCATGTCGATTTGAGTTGGTAGTACCTGTG 667
Db 613 GTTGGCATGTTTAACTTCACGTCATGTCGATTTGAGATTTGGTAGTACCTGTG 554

QY 668 GTTTCCTTGGCTGCTATAGTTGTAATA 695
Db 553 GTTTCCTTGGCTGCTATAGTTGTAATA 526

RESULT 11
AU236368
LOCUS
DEFINITION AU236368 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-93-K05 5',
647 bp mRNA linear EST 01-APR-2002
AU236368
Accession
AU236368
Version
AU236368.1 GI:19875537
Keywords
EST.
Source
Arabidopsis thaliana (thale cress)
Organism
Arabidopsis thaliana

REFERENCE
AUTHORS
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 647)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sekurai, T., Carinci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shingawa, A., Muramatsu, M., Hayashizaki, Y.,
and Shinzaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda PhC-1 vector (Carinci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for further
details.

FEATURES
source
1.647
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL14-93-K05"
/tissue_type="root"
/lab_host="DH10B"
/clone_lib="RAFL14"
/note="Site_1: BamHI; Site_2: SalI"

ORIGIN
Query Match 12.1%; Score 84; DB 9; Length 647;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 AGTTATTTTCTCCCGGAGTCGACACCTACCTCCTCCGCGCTTAACT 258
Db 1 AGTTATTTTCTCCCGGAGTCGACACCTACCTCCTCCGCGCTTAACT 60

QY 259 TACGTTCTCCGTCGTTACTCTGT 282
Db 61 TACGTTCTCCGTCGTTACTCTGT 84

RESULT 12
BZ558518

LOCUS
DEFINITION BZ558518 1872 bp DNA linear GSS 17-DEC-2002
pa98401.209.x1 pacs2-164 Pseudomonas aeruginosa genomic clone
pa98401.209, genomic survey sequence.

ACCESSION
BZ558518
VERSION
BZ558518.1 GI:27173068
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 1872)
REFERENCE
AUTHORS
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source
1.1872
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pa98401.209"
/clone_lib="pacs2-164"
/note="Clinical isolate 2-164 Whole genomic shotgun
library."

ORIGIN
Query Match 11.9%; Score 82.8; DB 28; Length 1872;
Best Local Similarity 46.2%; Pred. No. 0.00012;
Matches 304; Conservative 0; Mismatches 348; Indels 6; Gaps 1;

QY 2 ACACACAAACAAATTAAGCAGTCATGTCGATTAATTAATAGTGGAAACAAT 61
Db 527 AAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 586

QY 62 TAACTTAACGAGAAAGAGAAAGAGTCAAAATGAAACAAATCAACTGAATG 121
Db 587 AAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 646

QY 122 AAAATTTGAGTCAGATCGGAAACGAGCCGTTTAGAGCTTAATAGCTTCCTCA 181
Db 647 AAAAAAAAAAGGGGAGGAAAAAAAAAGGGGGGTTTGGGGGTTTGTGTGTGT 706

QY 182 TTTGCTCTCTCTCGTCACTTAATTTCTCTCTCGGAGTCCTGACTCACTCTCACT 241
Db 707 TTTTCT 766

QY 242 CTCGGCGCTTAACTTACGTCCTCCGCTTCTGTAAGTTTTCGCTTAGAC 301
Db 767 TGTGTGTT 826

QY 302 CTCGCATCCGCTCACGCGATGCTGTCGATTTCTCTTCTCTCGCTGAAAAA 361
Db 827 TTTTCT 886

QY 362 TTGCTTAATGTTCTCGATTTGAAAGTTTGTGCTAGAGTTACTTTTCCCTAAT 421
Db 887 TTGCTTAATGTTCTCGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 946

QY 422 TTATAGTCTTAGTAAGTAAGTAAGTACGCTCTTACGTTTGTGATTTGTGCTT 481
Db 947 TTTTATTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 1006

QY 482 TCACGTTTAGTCGATCGAGTAATTGACTGTAATAATCCTTCGTTTTGGTTT 541

Db 1007 TTTT... 1066
 QY 542 TGTTCATATAATC-----GATGATCTACCTTTGCTGATGTTGTTTGG 595
 Db 1067 TATTTGTAATATTTT... 1126
 QY 596 AGCCTAGCGCTGTGCTGTATTAACCTGACGTGATGTCGATTTGAGATTT 653
 Db 1127 TTTT... 1184

RESULT 13
 CDS00ETV 1068 bp DNA linear GSS 04-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
 DEFINITION BACR29P07 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL069846 GI:4949798

ACCESSION
 VERSION AL069846.1 GI:4949798
 KEYWORDS GSS.
 ORGANISM Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1068)
 Genoscope.

REFERENCE
 AUTHORS
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqrel@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

COMMENT
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazuo Ooegawa and
 Aaron Mammeter in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp, the same strain used for the BDGP's
 pl and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 location/Qualifiers

FEATURES
 source
 1.1068
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR29P07"
 /clone_lib="RPCI-98"
 /note="end : T7"

ORIGIN

Query Match 11.7%; Score 81.2; DB 29; Length 1068;
 Best Local Similarity 35.6%; Pred. No. 0.00031;
 Matches 245; Conservative 82; Mismatches 361; Indels 0; Gaps 0;
 QY 2 ACAAACAACAATAAGCAGTATGCTGACATTAATTAATGCGAACAT 61
 Db 248 AA 307
 QY 62 TAAGTTAGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121
 Db 308 AA 367
 QY 122 AAAATTGAGTCGAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAG 181
 Db 368 AA 427

QY 182 TTGCTCTCTTCGTCAGTTATTTTCTCCGAGAGCTCAGTACTGCTACT 241
 Db 428 MAAA 487
 QY 242 CTCGGCGCTTTAACTGATGCTTCGCGTTTACTGTGAAGTTTCGCTAGAC 301
 Db 488 TTTTTCMMMMMMMTTTTCTCMMMMMMMMMTTHTTTTTCCTCCCCCC 547
 QY 302 CTCGAGTCGCCACCGAGATGCTGTCGATTTCTCTTTCTTCGCTGAAAA 361
 Db 548 CCTTTTCTTCMCMMTTTCYCCCCCCCCCCCCCCCCCCCCCCCC 607
 QY 362 TTGCCCTAATGTTCTGATTCGAAGTTTGTGCTATGAGTTACTTTTCCGAT 421
 Db 608 TAAATTTTWTATTAATTTTTHATTTTWTATTAATAAAAAAAAAATTAAT 667
 QY 422 TTTATGTTCTTAGTACGATACGATCGCTCTTACTGTTTGTTCATTTGTCCT 481
 Db 668 TTTTWTATAAAAAAAAAAAAAAAAAMTTTTTHATTAATTTTATTAATTTGAMAMKT 727
 QY 482 TCACGCTTAGTCGATCGAGATTTGACTGTGMAAAATCCTTCGTTTGTGTTT 541
 Db 728 GTMAATTTTAAAAAATAATTTTCTTTTATTAATTAATAATAATAATAATA 787
 QY 542 TGTTCATATAATCGATGATGATCTACCTTTTGTGCTTGTGATTTTGTGACCTA 601
 Db 788 ATTTTCTTTTGTGMAAATAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 847
 QY 602 TCGCTTTGCTGTTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 661
 Db 848 KGTTRKGTTRKGTTRKGTTRKGTTRKGTTRKGTTRKGTTRKGTTRKGTTRK 907
 QY 662 CTGCGGCTTTCTTGAGGCTTAGGTT 689
 Db 908 GTGTGKKTGTGTGTGKKTGTGTGKKT 935

RESULT 14
 BH791768/c 81 bp DNA linear GSS 02-APR-2002
 LOCUS SALK_061198.45.85.x Arabidopsis thaliana TDNA insertion lines
 DEFINITION Arabidopsis thaliana genomic clone SALK_061198.45.85.x, genomic
 survey sequence.

ACCESSION
 VERSION BH791768.1 GI:19885974
 KEYWORDS GSS.
 ORGANISM Arabidopsis thaliana (chale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
 Gadinb,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
 Shum,P., Zimmerman,J. and Ecker,J.R.

TITLE
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 JOURNAL Unpublished (2001)
 COMMENT Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu

This is single pass sequence recovered from the left border of
 TDNA.
 Class: TDNA tagged.

FEATURES
 source
 1.81
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia 0"

/db_xref="taxon:3702"
/clone="SALK_061198.45.85.x"
/clone_1ib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html"

ORIGIN

Query Match 11.7%; Score 81; DB 28; Length 81;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 AGCTTCCTAATTTGCTCTCTTCCTGTCAGTTATTTTCTCCGAGTCCTGACTCAG 231
DB 81 AGCTTCCTAATTTGCTCTCTTCCTGTCAGTTATTTTCTCCGAGTCCTGACTCAG 22

QY 232 TACTCTACCTCTCCGGGCTT 252
DB 21 TACTCTACCTCTCCGGGCTT 1

RESULT 15 1201 bp mRNA linear EST 08-MAY-2003
AL514657/c LOCUS AL514657 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION CL08B0022E09 3-PRIME, mRNA sequence.
AL514657
AL514657.2 GI:30464542

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:1278151.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6668.f. For
more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CL08B0022E09&pkcluster=6668.f>. Contact :
Feng Liang Email: liang@lifetech.com URL: <http://fulllength.invitrogen.com/InvitrogenCorporation1600>
Faraday Avenue Genoscope sequence ID: CL08B0022E09FP1.
Location/Qualifiers
1. 1201

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL08B0022E09"
/tissue_type="NEUROBLASTOMA"
/clone_1ib="Homo sapiens NEUROBLASTOMA"
/note="Vector: PCWSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime and enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the PCWSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 11.6%; Score 80.4; DB 9; Length 1201;
Best Local Similarity 40.0%; Pred. No. 0.00039;
Matches 220; Conservative 50; Mismatches 280; Indels 0; Gaps 0;

QY 142 GGAAAAAGAGCGCGTTTAGAGCTTAATAGCTTCCCTCATTTGTCTTCTGTCAGT 201

DB 842 GGAAACCGDKTCCSCMCWAAAAAAMWRARACHTTWTGTTTCRVTGGGGR 783
QY 202 TTAATTTTCTCCCGGAGATCCGACTCAGTCTGACTCTCCGCGCTTAATAC 261
DB 782 GDAATTTTGTGKRAATTTTTCCTTTTCCCAAGGKGKATGDKGAAR 723
QY 262 GTTTCGCGTGTACTCTGTAAGTTTTCGCTTAGAGCTTCGATCCGAT 321
DB 722 GGVCCCCCHTHNMGSCCCCMKTTTGTGSKSCCTGAGNCCCCCTMSCCCMTWT 663
QY 322 GAATTCGTGCTGATTTCTTTTCTTCGCTGAAAAATGCGCTTAATGTTGCAAT 381
DB 662 TTTTTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 603
QY 382 TCGAAGTTTGTGCTAGAGGTACTTTTCCATATTTATAGTTCTTAAGTAAG 441
DB 602 TGTATTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 543
QY 442 ATACGTGCTTACTGTTTTTGTTCATTTTGTGTTGCTTACGCTTAAGTCGATC 501
DB 542 TTTTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 483
QY 502 GAGTATTTGACTGTGAAAAATCCTTCGTTTTTGTGTTTTGTTCAATAATCGAAT 561
DB 482 TTTTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 423
QY 562 GATCTACCTTTGTGCTTGTGATGTTTGTGAGCTATGCGTTGTGCTGTATA 621
DB 422 TKTGTGTTT 363
QY 622 ACTTCAGTTCAATGTTGATTTTGTAGATTTTGTGATGATGCTTCTTGTGGC 681
DB 362 GCGCCCTT 303
QY 682 TATAGTTGT 691
DB 302 TTTTTCCTT 293

Search completed: April 8, 2004, 16:24:28
Job time: 1220.46 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 7, 2004, 19:34:49 ; Search time 35.059 Seconds
(without alignments)
11001.196 Million cell updates/sec

Title: US-10-030-829-1_COPY_1_695
Perfect score: 695
Sequence: 1 gcaacaacaacaataa.....ggtgctatagtgtaaa 695

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/6CTUS COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85.6	12.3	7218	1	US-08-232-463-14 Sequence 14, Appl
2	52.6	7.6	7286	3	US-09-331-581-3 Sequence 3, Appl1
3	52.6	7.6	7938	4	US-09-331-581-14 Sequence 14, Appl
4	50.6	7.3	396	4	US-09-640-173-53 Sequence 53, Appl
5	50.6	7.3	396	4	US-09-713-550-53 Sequence 53, Appl
6	50.6	7.3	10619	4	US-10-204-708-4 Sequence 4, Appl1
7	50.2	7.2	5562	4	US-10-204-708-63 Sequence 63, Appl
8	49.4	7.1	19124	2	US-08-487-826B-13 Sequence 13, Appl
9	47.8	6.9	2447	2	US-09-014-969-14 Sequence 14, Appl
10	47.4	6.8	5666	4	US-10-204-708-29 Sequence 29, Appl
11	47.4	6.8	8607	4	US-10-204-708-72 Sequence 72, Appl
12	46.8	6.7	19233	4	US-10-204-708-45 Sequence 45, Appl
13	46.8	6.7	19233	4	US-10-204-708-45 Sequence 45, Appl
14	45.6	6.6	7218	1	US-08-232-463-14 Sequence 14, Appl
15	45.6	6.6	289	3	US-09-007-005-17 Sequence 17, Appl
16	45.6	6.6	11049	4	US-10-204-708-23 Sequence 23, Appl
17	45.4	6.5	11049	4	US-10-204-708-23 Sequence 23, Appl
18	44.2	6.4	1493	1	US-08-340-820-24 Sequence 24, Appl
19	44.2	6.4	1493	1	US-08-593-535-24 Sequence 24, Appl
20	44.2	6.3	240	1	US-08-628-417-6 Sequence 6, Appl1
21	43.6	6.3	2394	4	US-09-800-729-33 Sequence 33, Appl
22	43.6	6.3	359	4	US-09-621-976-16008 Sequence 16008, A
23	43.6	6.3	2915	4	US-09-336-115C-5 Sequence 5, Appl1
24	43.6	6.2	5501	4	US-10-204-708-38 Sequence 38, Appl
25	43.4	6.2	140	1	US-08-628-417-5 Sequence 5, Appl1
26	43.4	6.2	359	4	US-09-621-976-16019 Sequence 16019, A
27	43.4	6.2	362	4	US-09-621-976-16010 Sequence 16010, A

28	43.4	6.2	365	4	US-09-621-976-16042 Sequence 16042, A
29	43.4	6.2	5455	4	US-10-204-708-33 Sequence 33, Appl
30	43.4	6.2	11015	4	US-10-204-708-55 Sequence 55, Appl
31	43	6.2	578	4	US-09-602-877A-95 Sequence 95, Appl
32	43	6.2	8607	4	US-10-204-708-72 Sequence 72, Appl
33	43	6.2	8961	4	US-10-204-708-80 Sequence 80, Appl
34	43	6.2	19124	2	US-08-487-826B-13 Sequence 13, Appl
35	42.8	6.2	1114	4	US-09-152-060-41 Sequence 41, Appl
36	42.8	6.2	11050	4	US-10-204-708-85 Sequence 85, Appl
37	42.8	6.2	11131	4	US-10-204-708-27 Sequence 27, Appl
38	42.6	6.1	240	1	US-08-628-417-6 Sequence 6, Appl1
39	42.6	6.1	371	4	US-09-621-976-16048 Sequence 16048, A
40	42.6	6.1	10144	4	US-10-204-708-94 Sequence 94, Appl
41	42.4	6.1	270	2	US-08-520-678A-30 Sequence 30, Appl
42	42.4	6.1	270	3	US-08-897-126-30 Sequence 30, Appl
43	42.2	6.1	5844	4	US-10-204-708-89 Sequence 89, Appl
44	42.2	6.1	6866	4	US-10-204-708-20 Sequence 20, Appl
45	42	6.0	357	4	US-09-621-976-16058 Sequence 16058, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
CLASSIFICATION: 435
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT29PT-F18
US-08-232-463-14
Query Match 12.3%; Score 85.6; DB 1; Length 7218;

Query Match 7.3%; Score 50.6; DB 4; Length 396;
Best Local Similarity 49.8%; Pred. No. 0.00049;
Matches 119; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 389 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 448
DB 18 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 77

QY 449 CGCTTACGCTTTGCTTCAATTTGCTTGCCTTACCGCTTAGTGCGCTAGCGAGTAT 508
DB 78 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 137

QY 509 TTGACTGTGAAAAATCCTTCGTTTGTGTTGCTTCAATAAATCGATTGATCTAC 568
DB 138 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 197

QY 569 CTTTGTGCTTGAATGTTGTTTGGAGCTTATGCGTGTGCTGTATTAATCTCA 627
DB 198 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 256

RESULT 5

US-09-713-550-53
Sequence 53, Application US/09713550
Patent No. 6617109

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Stoik, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
FILE REFERENCE: 210121.484C4
CURRENT APPLICATION NUMBER: US/09/713,550
CURRENT FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 205
SOFTWARE: FaSTSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(396)
OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

Query Match 7.3%; Score 50.6; DB 4; Length 396;
Best Local Similarity 49.8%; Pred. No. 0.00049;
Matches 119; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 389 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 448
DB 18 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 77

QY 449 CGCTTACGCTTTGCTTCAATTTGCTTGCCTTACCGCTTAGTGCGCTAGCGAGTAT 508
DB 78 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 137

QY 509 TTGACTGTGAAAAATCCTTCGTTTGTGTTGCTTCAATAAATCGATTGATCTAC 568
DB 138 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 197

QY 569 CTTTGTGCTTGAATGTTGTTTGGAGCTTATGCGTGTGCTGTATTAATCTCA 627
DB 198 TTTTGTGCTAAGGCTTCTTTTCCCTATATTAAGTCTTAGGTAACATACCTG 256

RESULT 6

US-10-204-708-4
Sequence 4, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: Olek, Alexander
APPLICANT: Piepenbrock, Christian

APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 4
LENGTH: 10619
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-4

Query Match 7.3%; Score 50.6; DB 4; Length 10619;
Best Local Similarity 49.8%; Pred. No. 0.0017;
Matches 128; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 404 TTACTTTTCCCTATATTTATTAAGTCTTAGGTAACATACCTGCTTACTGTTTTT 463
DB 9137 TTAATTTATTTGTAATATTAATGATTTATTTATTTATTTATTTATTTATTTT 9196

QY 464 GTTCAATTTGTTGCTTTCACCGTTTATGCTGATCGAGATTGACTGAAAAAT 523
DB 9197 TTTCGTTTGTGTTTTTTTTTTTTTTTTTTCGTTTGAATGATTTATTTGTTATAT 9256

QY 524 CCTGCTTTTGTGTTTGTTCATATAATGATGATACCTTTGCTTGTAT 583
DB 9257 AGTTTTTTTGTGTAATGATGATGATTTGTTTGTGATTTTATTTTATGAGAA 9316

QY 584 GTTGTGTTTGTGAGCTTATGCTTGTGCTTGTATTAATCACTTCAGTGTGAT 643
DB 9317 GTTGTGTTTGTGCTTGTGTTATTTGATGATGATGATTTATTTATTTATTTAGTTT 9376

QY 644 TTGAGATTTGTTAGTG 660
DB 9377 TGGTATTTGTTAGTG 9393

RESULT 7

US-10-204-708-63
Sequence 63, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: Olek, Alexander
APPLICANT: Piepenbrock, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 98
SEQ ID NO: 63
LENGTH: 5562
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-63

Query Match 7.2%; Score 50.2; DB 4; Length 5562;
Best Local Similarity 48.1%; Pred. No. 0.0017;
Matches 142; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 390 TTTTGCTATGAGGCTACTTTTCCCTAATTTTAACTTCTAGTAACGACTCTGC 449
DB 1394 TATTTATTAATTAATTTTATTTTTCGTTAGGAAGATTTTTCGATTAATTTAT 1453
QY 450 GTCTTACTGTTTGTCTCAATTTGTGCTTACCGTTAGCTGCTGAGAGTAT 509
DB 1454 ATTTTATTTTATTTAGTTAGTTATTTTATTTTATTTTATTTTATTTTATTTT 1513
QY 510 TGAAGTGAAGAAATCTGCTTTTGTGTTTGTCTTCAATTAATCGATGATCTAC 569
DB 1514 TGTATCTAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1573
QY 570 TTTTGCTTGAATTTTGTGTTTGTGAGCCTATGCTGTTGCTTGAATTAATCTCAG 629
DB 1574 TTTTTCGTTTGAAGAAATTTATTTATTTTATTTTATTTTATTTTATTTTATTT 1633
QY 630 TTCATGCTGATTTTGAATTTTGTGATGCTGCTGCTTCTTGTGCTAT 684
DB 1634 GATAGGTATTTTATTTTGAATTTTGTATTAATTAATTTTATTTATTTATTTAT 1688

RESULT 8
US-08-487-826B-13/C
Sequence 13, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas B.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:

LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 7.1%; Score 49.4; DB 2; Length 19124;
Best Local Similarity 47.9%; Pred. No. 0.0042;
Matches 172; Conservative 0; Mismatches 186; Indels 1; Gaps 1;

QY 318 GCATGATCTGCGTCGATTCCTCTTTCTTCGCGTGAAGAAATTCCTAATGTTCTC 377
DB 15976 GCATGATTAATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15917
QY 378 GATTCGAAGTTTGTGCTAATGCTTACCTTTTCCCTAATTTTATTAATGTTAGGT 437
DB 15916 TCATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15857
QY 438 AACGATCTGCTGCTTACTGTTTGTTCATTTTGTGCTTACCGTTAGCTGCT 497
DB 15856 TATTTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15797
QY 498 GATCGAGATTTGACTGAGAAATCTGCTTTTGTGTTTGTTCATTAATTAATCG 557
DB 15796 TATGATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15737
QY 558 G-ATTGATCTACCTTTTGTGCTTGAATTTGATTTTGTGAGCCATGCTGTTGCTTG 616
DB 15736 TATTTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15677
QY 617 TATTAATCTGAGTCACTGCTGATTTTGAATTTTGTGATGCTGAGGTTCTT 675
DB 15676 AATGTTTATTTTCTCTTGTGTTTATTTTATTTTATTTTATTTTATTTTATTTAT 15618

RESULT 9

US-09-014-969-14/C
Sequence 14, Application US/09014969
Patent No. 5965397

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racine, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2447 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-014-969-14

Query Match 6.9%; Score 47.8; DB 2; Length 2447;
Best Local Similarity 47.2%; Pred. No. 0.0051;
Matches 142; Conservative 1; Mismatches 158; Indels 0; Gaps 0;

QY 390 TTTTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGTACGATACCTGC 449
DB 2447 TTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGTACGATACCTGC 2388
QY 450 GCTTACGTTTGTGCTATTTGTTGCTTACCGTTAGCTGATCGAGATTT 509
DB 2387 TTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGTACGATACCTGC 2328
QY 510 TGACGTGAAAAATCCCTGTTTGTGTTTGTTCATATAAATCGATTGATCTACC 569
DB 2327 TTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGTACGATACCTGC 2268
QY 570 TTTTGTGCTATGGGTTACTTTTGTGCTTACCGTTAGCTGATCGAGATTT 629
DB 2267 TTTTGTGCTATGGGTTACTTTTGTGCTTACCGTTAGCTGATCGAGATTT 2208
QY 630 TTCATGCTGATTTGATTTTGGATGAGTGGGTTTGTGCTTATAGTT 689
DB 2207 ATCAATACACATCCAGTATGTGAACCATATATACATATATACGCCATTATTA 2148
QY 690 G 690
DB 2147 G 2147

RESULT 10
US-10-204-708-29
; Sequence 29, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204, 708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 29
; LENGTH: 5666
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-29

Query Match 6.8%; Score 47.4; DB 4; Length 5666;
Best Local Similarity 49.2%; Pred. No. 0.0088;
Matches 153; Conservative 0; Mismatches 156; Indels 2; Gaps 1;

QY 384 GAAGTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGTACGAT 443
DB 3358 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3417
QY 444 ACCGCGCTTACGTTTGTGCTTACCGTTAGCTGATCGAGATTT 503
DB 3418 TTTTACGTTATTTTGTGCTTACCGTTAGCTGATCGAGATTT 3477
QY 504 AGATTTGACGTGAAAAATCCCTGTTTGTGTTTGTTCATATAAATCGAT--T 561
DB 3478 TTTTGTGCTATGGGTTACTTTTGTGCTTACCGTTAGCTGATCGAGATTT 3537
QY 562 GATCTACCTTTGTGCTTATGTTTGTGCTTACCGTTAGCTGATCGAGATTT 621
DB 3538 GGTGTTGCTTGTGTTTGTGCTTACCGTTAGCTGATCGAGATTT 3597
QY 622 ACTTACGTTTGTGCTTATGATTTGTGCTTACCGTTAGCTGATCGAGATTT 681
DB 3598 TGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3657
QY 682 TATAGTTGTA 692
DB 3658 TATGCTTATA 3668

RESULT 11
US-10-204-708-72
; Sequence 72, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204, 708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 72
; LENGTH: 8607
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-72

Query Match 6.8%; Score 47.4; DB 4; Length 8607;
Best Local Similarity 46.7%; Pred. No. 0.01;
Matches 150; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 374 TCTGATTCGAAGGTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTT 433
DB 6966 TATGTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 7025
QY 434 AGTACGATACCTGCGTTTACGTTTGTGCTTACCGTTAGCT 493
DB 7026 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 7085

QY 494 CCGTATCGAGTATTTGACGTGAAAAATCCTTCTTTTGGTTTGGTTCATATA 553
 DB 7086 TTTTGGTTTATTTTGGTTTATATATTTTATAGTGGTTTGGTTCAGAA 7145
 QY 554 ATCGGATTAATCTACCTTTTGGCTTGAATTTTGGTTCAGCTTATGCGC 613
 DB 7146 ATAAAGAGGTTTATAGTGTGTTTGTGTTTATTTAGTTTATAGAAAGTTT 7205
 QY 614 TTGTATTAATCTACGCTTCAATGCTGATTTTGAATTTGGTATGATCTGGTTCT 673
 DB 7206 AGAATTAGTATTTTGTGATAGTATTTGTGACGTGAAAGGATATGATTTTA 7265
 QY 674 TTGGTGGCTATAGTTGTATA 694
 DB 7266 AGTATGTTTATTTTGGTATA 7286

RESULT 12

US-10-204-708-79
 ; Sequence 79, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; PRIOR FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 79
 ; LENGTH: 8961
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; NAME/KEY: unsure
 ; LOCATION: (3866)
 ; OTHER INFORMATION: n is a or g or c or t
 ; US-10-204-708-79

Query Match 6.8%; Score 47; DB 4; Length 8961;

Best Local Similarity 48.9%; Pred. No. 0.013;

Matches 155; Conservative 0; Mismatches 160; Indels 2; Gaps 1;

QY 341 TCTTTTCTTCGCTGAAAAATGCGCTTAATGTTTCGATTCGAAAGTTTGGCTAT 400
 DB 6013 TATTTGTTGATGCGGGAATGTTGTTTATTTTAAAGATGATGTTAGTTT 6072
 QY 401 GGGTACTTTTCCCATATTTTATAGTCTTAGAGTAACGATACGCGCTTACTGT 460
 DB 6073 TAAGTAAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6132
 QY 461 TTGTTCATTTTGTGTCCTTCAACGTTAGTCCGTAACGAGATTTGACTGGAA 520
 DB 6133 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6190
 QY 521 AATCCTTGGTTTGTGTTTGTTCATATAAAGCATGATCTACTTTGTGCTT 580
 DB 6191 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 6250

QY 581 GATTTTGTTTTGAAGCCTATGCGTTGTGCTTGTATATACATCACTTCATGTG 640
 DB 6251 ATTATGATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6310
 QY 641 ATTATGATTTTGTGA 657
 DB 6311 TGGAGAGATGTTTATTA 6327

RESULT 13

US-10-204-708-45
 ; Sequence 45, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; PRIOR FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 45
 ; LENGTH: 19233
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; NAME/KEY: unsure
 ; LOCATION: (76, 178, 179, 273, 586, 648, 651, 920, 1014, 1173, 1197, 1228)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1264, 1267, 1559, 1607, 1677, 1703, 1742, 1848, 1850, 2033)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (2168, 2254, 2257, 2280, 2291, 2723, 3018, 3080, 4205, 4209)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (4216, 4225, 4239, 4246, 4254, 4265, 4270, 4274, 4288, 4294, 4295)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (4298, 4363, 4610, 4627, 4675, 4693, 4698, 4738, 4777, 4781)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (4787, 4825, 4829, 4838, 4855, 5462, 5494, 5496, 5500, 5674)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (5690, 5697, 5705, 5715, 5890, 5910, 5913, 6027, 6080, 6109)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (6244, 6288, 6776, 6873, 6891, 6980, 7028, 7656, 7876, 8030)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure

LOCATION: (8059, 8075, 8632, 8636, 9700, 9717, 9815, 9818, 9867, 9879)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (10005, 10075, 10150, 10159, 10165, 10356, 11123, 11166, 11360)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (11703, 11933, 12050, 12340, 12981, 13564, 13764, 13765, 13787)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (13846, 14168, 16566, 17006, 17217, 17650, 17653, 17656, 17898)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (18130, 18143, 18151, 18188, 18498, 18549, 18936, 18946, 19165)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (19200)
OTHER INFORMATION: n is a or g or c or t
US-10-204-708-45

Query Match 6.7%; Score 46.8; DB 4; Length 19233;
Best Local Similarity 48.5%; Pred. No. 0.02;
Matches 129; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 360 AATTGCCCTTAAGTCTCGATTCGAAAGGTTTGGCTAGTGGCTACTTTTCCCTAT 419
DB 11391 AATTAAATGTTTGTGGAGATATGAAATTTTATTAACGTAATATTTTGTGTT 11450
QY 420 ATTCTAGTCTTAGGTAACGATACCTGCTTACTGTTTGTTCATTTGTTGTC 479
DB 11451 TTTTATTTTATTTTATTTTATTTAGTAATGTTGTTTGTATTTTGTAGGTT 11510
QY 480 TTTCACGCTTATCGCTGATCGAGATTTGACTGGAATAATCCCTCGTTTGTGTT 539
DB 11511 AGTTTGTGTTTATTAAGATTTATATATGAATTAATATATTTATTTTGTGT 11570
QY 540 TTTGTTTCAATATAATCGATGATGATCTACCTTTGTGCTTGTGATGTTTGTGAGCC 599
DB 11571 TTGGTTTTTTTGTAGTATTTTGTGTTTGTGTTTGTAGAGTGGGTTTGTATTTAGGTT 11630
QY 600 TATGCTGTGCTGTGCTTATACCT 625
DB 11631 GGAGTGTAGCATATATGTTGTTT 11656

RESULT 14

US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Jardner
STREET: 1800 diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pC-Fls
US-08-232-463-14

Query Match 6.6%; Score 46; DB 1; Length 7218;
Best Local Similarity 5.6%; Pred. No. 0.022;
Matches 10; Conservative 115; Mismatches 55; Indels 0; Gaps 0;

QY 1 GACAAACAACAATAATGACAGTCACTGTCGATCAATAATTAATGAGGGAACA 60
DB 1231 RRR 1172
QY 61 TTAAGTAAAGCAAGAAAAAGAAAAAGTACAAAAATGAAAAACAATCAACTGAAT 120
DB 1171 RRR 1112
QY 121 GAAATTTGAGCGCAGATCGAATAACGAGCGCTTTAGAGCTAATAAGCTCTC 180
DB 1111 RRR 1052

RESULT 15

US-09-007-005-17/C
Sequence 17, Application US/09007005B
Patent No. 6258358
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rih
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
FILE REFERENCE: 00786/350003
CURRENT APPLICATION NUMBER: US/09/007,005B
CURRENT FILING DATE: 1998-01-14
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
NAME/KEY: misc_feature
LOCATION: (1)...(289)
OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match 6.6%; Score 45.6; DB 3; Length 289;
 Best Local Similarity 10.4%; Pred. No. 0.0085;
 Matches 27; Conservative 105; Mismatches 128; Indels 0; Gaps 0;

```

QY 175 TTCCCTATTGCTCTCTTCGTCAGTTTATTTCTTCCTCCGAGTCCTGACTCAGTAC 234
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 275 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 TCTCAGCTCTCCGCGCTTAACTTAAGTTCCTCCGCTTACTCTGTAAGTTTCTGCC 294
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 215 YNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYN 156
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 295 TTAGAGCTCTCCGATCCGCTCAGCGCATGCTGCTGATTCCTCTTTCTTCGCT 354
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 155 YNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYN 96
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 355 GGAATAATGCGCTAATGTTCTCGATTTCGAAGTTTGTGCTATGGTTACTTTTTC 414
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 95 YNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYN 36
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 415 CCTATATTTTATAGTCTTA 434
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 35 YTAAYAAAYTAAGYTAAYA 16
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  
```

Search completed: April 8, 2004, 16:29:39
 Job time : 36.059 secs

QY	260	ACGTTCTCCGCGTGTACTCTGTAAAGTTTCTGCCTTAAGACCTCCGATCGCTACCGG	319
Db	971629	TT	971688
QY	320	ATGCATCTGTCGATTCGATTCCTCTTTCTCGCGAAAAATGCCCAATGTCCTCGA	379
Db	971689	TT	971748
QY	380	TTTCCAGGTTTTGTGCTATGGGTACTTTTTCCTATATTATATAGTCTTAGCTAA	439
Db	971749	TTTTAATTT	971808
QY	440	CGATACCTGCGCTACTGTTTTTGTCATTTTTGTGTGCTTACCGTTAGTCGTGA	499
Db	971809	TTAATTTTTTTTTTTTTTT	971868
QY	500	TCGAGATATTGACGTGAAAAATCCTGTTTTTGTGTTTCATATATAATCGGA	559
Db	971869	TTTTTTTTTTTTTTTTTTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	971928
QY	560	TTGATCTACCTTTGTGCTTTCATGTTTGTTTTTGAGCCCTAAGCSTGTGGCTGTGA	619
Db	971929	TTTTTTTTTTTTTTAATTT	971988
QY	620	TAACTCAAGTTCACTGTGAGATTTTGAGATTTTGGTAGACTGCGGTTCTTTGG	677
Db	971989	TTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGATAGAGATTTGTTTTGTAG	972046

```

RESULT 2
US-10-311-455-1670
: Sequence 1670, Application US/10311455
: Publication NO. US20030143606A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIEPENBROCK, Christian
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
: TITLE OF INVENTION: Cytosine methylation
: FILE REFERENCE: 5013.1014
: CURRENT APPLICATION NUMBER: US/10/311,455
: CURRENT FILING DATE: 2002-12-16
: PRIOR APPLICATION NUMBER: PCT/EP01/07537
: PRIOR FILING DATE: 2001-07-02
: PRIOR APPLICATION NUMBER: DE 10032529.7
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: DE 10043826.1
: PRIOR FILING DATE: 2000-09-01
: NUMBER OF SEQ ID NOS: 2424
: SEQ ID NO 1670
: LENGTH: 6668
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 1936
: OTHER INFORMATION: n is a or g or c or t
: US-10-311-455-1670

```

	Query Match	9.9%	Score 68.8	DB 14	Length 6668
	Best Local Similarity	45.8%	Pred. No. 7.8e-06		
	Matches	238	Conservative	0	Mismatches 282
				Indels	0
				Gaps	0
Qy	156	GTTTTAAGCTTAATAAGCTTCCTCATTTGCTCTTCTTCGACATTATTTCTCTCTC	215		
Db	2742	GTTTTGTGTTTGT	2801		
Qy	216	CGAGTCCTGACTCACTCACTCCGAGCGTTAAACTAAGTTCCTCGTGTTT	275		
Db	2802	TTTTTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	2861		

[illegible]

```

RESULT 3
US-10-239-676-52
Sequence 52, Application US/10239676
Publication No. US20030082609A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
FILE REFERENCE: 5013.1003
CURRENT APPLICATION NUMBER: US/10/239,676
CURRENT FILING DATE: 2002-09-24
PRIOR APPLICATION NUMBER: PCT/EP01/03968
DE 10019058.8
DE 10019173.8
DE 10032529.7
DE 10043826.1
PRIOR FILING DATE: 2001-04-06
2000-04-06
2000-04-07
2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 52
LENGTH: 9539
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-52

```

Query Match	9.64;	Score 66.6;	DB 14;	Length 9539;
Best Local Similarity	46.94;	Pred. No. 3,1e-05;		
Matches 207;	Conservative	0;	Mismatches 234;	Indels 0;
			Gaps	0;
QY	251	TTTAACTTACGTTCTTCGCGTTTACTCGTAAGTTTTCGCTTACAGCGCTCCGATCG	310	
Db	202	TTTTTTTTTTTGATTTTTTTCGTTTATTTTTTGGTGTTTTTTTTTTTATTTTTTTTTTTC	261	
QY	311	CCTACCGCAGCATTCGTGCTGATTTCTTTTTCTTCGCTGGAATAATGCGCTAA	370	
Db	262	GTTTTTTTTTTTTTGGCGTTGGTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTAAAT	321	
QY	371	TGTTTCGATTTGGAAGGTTTTTGTGCTAAGGTTACTTTTTTCCCAATATTTATAGTT	430	

[illegible]

RESULT 4

```

; Sequence 54 Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BEBLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
; TITLE OF INVENTION: With DNA Transcription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE_10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE_10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE_10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE_10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 54
; LENGTH: 9539
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-240-453--54

```

Query Match	9.6%	Score 66.6;	DB 14;	Length 9539;
Best Local Similarity	46.9%	Pred. No. 3.1e-05;		
Matches 207; Conservative	0;	Mismatches 234;	Indels 0;	Gaps 0

Qy	251	TTTAACTACGATTCCTCCGTCGTTACTCTGTAAAGTTTCTGGCTAAGAGCCCGCATCG	310
Db	202	TTTTTTTTTTGGATTTTTTTCGTTTTATTTTGTGGTTTTTTTTTTTATTTTTTTTTTC	261
Qy	311	CCTACCGCATGCATCTGTGCTCGATTTCTCTTTTCTCTGGTGAATAATGGCCCTAA	370
Db	262	GTTTTTTTTTTTTTGTCTGTTTGTTTTTTTTTTTTTTTTTTTTTTGTGTTTTTTTATTT	321
Qy	371	TGTTCTCGATTTGGAAGTTTTGTGCTAAGGGTAACTTTTTCCCTATATTTTAAAGTT	430
Db	322	TTAATTTGTTTTATTTTTTTTTTTTTTTTTTTTAAATTTTTTTTTTTTTTTTTTTTTT	381

Oy	431	CTTAGGTAACATACCTCGCTCACTGTTTGTTCATTTGTGTGCTTCACCGTT	490
Db	382	TTTTTTTGTGTTTTTTTCGTTTTTTTTTTTAAATTTTTTTTGTTTTTTTTTTT	441
Oy	491	AGTCGCTGATCGAGATATTGACCTGTGAAAAATCCTCGTTTTGTGTTTCATA	550
Db	442	ATTT	501
Oy	551	TAAATCGATTGATCTACCTTTTGCTTGATGTTGTTTTTTAGCCATAGCGTTGT	610
Db	502	TTTTTTTTTTTTTTGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	561
Oy	611	GCGTGTATAACTTCACGTCATGTCGATTTTGAGATTTGGTAGTGACTGCGGTT	670
Db	562	TTTAAATAATTAATTTTTTTT	621
Oy	671	TCATTGGGCGATAGTGT	691
Db	622	TTATTTAATAATTTTTTTT	642

RESULT 5

```

: Sequence 57885, Application US/10424599
: Publication No. US20040031072A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovall David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 57885
: LENGTH: 673
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(673)
: OTHER INFORMATION: unsure at all n locations
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT3847_23282C.1
: US-10-424-599-57885

```

Query Match	9.5%;	Score 66.2;	DB 12;	Length 673;
Best Local Similarity	47.0%;	Pred. No. 1.1e-05;		
Matches 170; Conservative	0;	Mismatches 192;	Indels 0;	Gaps 0;

[illegible]

QY 630 TTCATGCTGATTTTGGATTTTGGTAGACCTGCTGCTTTCTTTGGTGGCTATAGGT 689
 DB 480 TTTTCTTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 539
 QY 690 GT 691
 DB 540 AT 541

RESULT 6
 US-10-311-455-1931
 ; Sequence 1931, Application US/10311455
 ; Publication No. US20030143606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting
 ; TITLE OF INVENTION: Cytosine methylation
 ; FILE REFERENCE: 5013.1014
 ; CURRENT FILING DATE: 2002-12-16
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537
 ; PRIOR FILING DATE: 2001-07-02
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 2424
 ; SEQ ID NO 1931
 ; LENGTH: 14006
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; NAME/KEY: unsure
 ; LOCATION: 8289, 8310, 8313
 ; OTHER INFORMATION: n is a or g or c or t
 ; US-10-311-455-1931

Query Match 9.5%; Score 66; DB 14; Length 14006;
 Best Local Similarity 45.3%; Pred. No. 5.2e-05;
 Matches 240; Conservative 0; Mismatches 290; Indels 0; Gaps 0;
 QY 157 TTTTACGCTTAAATGCTTCTCATTTGCTCTTCTTCTGCTAGTTATTTCTCTCC 216
 DB 1974 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2033
 QY 217 GGAGTCTGACCTAGCTCTGCTGCGGCTTAACTAGTCTGCTGCTTAA 276
 DB 2034 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2093
 QY 277 CTCTGTAAGTTCTGCTAGAGCTCGATCGCTCAACGATCTGCTCGA 336
 DB 2094 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2153
 QY 337 TTTCTCTTTTCTGCTGAAAAATGCTTATGTTCTTCTTCTTCTTCTTCTTCTTCTT 396
 DB 2154 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2213
 QY 397 CTATGCTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 456
 DB 2214 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2273
 QY 457 TGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 516
 DB 2274 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2333
 QY 517 GAAAAATCTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 576
 DB 2334 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2393

QY 577 CTATGCTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 636
 DB 2394 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2453
 QY 637 GTGATTTTGAATTTTGTAGTACCTGCTGCTTTCTTTGGTGGCTATAG 686
 DB 2454 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2503

RESULT 7
 US-10-311-455-2147
 ; Sequence 2147, Application US/10311455
 ; Publication No. US20030143606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting
 ; TITLE OF INVENTION: Cytosine methylation
 ; FILE REFERENCE: 5013.1014
 ; CURRENT FILING DATE: 2002-12-16
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537
 ; PRIOR FILING DATE: 2001-07-02
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 2424
 ; SEQ ID NO 2147
 ; LENGTH: 113515
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; US-10-311-455-2147

Query Match 9.4%; Score 65.6; DB 14; Length 113515;
 Best Local Similarity 48.6%; Pred. No. 0.00018;
 Matches 179; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
 QY 325 TTTCTGCTCATTTCTCTTCTTCTGCTGAAAAATGCTTATGTTCTGATTTCTG 384
 DB 51453 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 51512
 QY 385 AAGGTTTGTGCTAGGCTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 444
 DB 51513 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 51572
 QY 445 CCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 51632
 DB 51573 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 51632
 QY 505 GATTTGACCTGAAAAATCTGCTTTTGTGTTTGTGTTTCTTCTTCTTCTTCTTCTTCTT 564
 DB 51633 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 51692
 QY 565 CTACCTTTTGTGCTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 624
 DB 51693 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 51752
 QY 625 TCAGCTCATGCTGATTTTGTAGATTTGTAGTACCTGCTGCTTCTTGTGCTGCTAT 684
 DB 51753 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 51812
 QY 685 AGGTTGTA 692
 DB 51813 AAGTTATA 51820

RESULT 8
 US-10-311-455-394
 ; Sequence 394, Application US/10311455

Query Match	9.3%;	Score 64.6;	DB 14;	Length 6668;
Best Local Similarity	49.3%;	Pred. No. 7.7e-05;		
Matches 169; Conservative	0;	Mismatches 174;	Indels 0;	Gaps 0;

```

      565  CTAAGTTTGGCTTGGATGTTGTTTGTGAGACCCAGACCGTTGGCTGTAACT 624
      Db    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      17819 TTTTGTGTTTTTTTGGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 17878
      QY     | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      625   TCAGCTTATGTTGATTTTGAGATTTGGTAGTGTGGGTTCTTGGT 678
      Db     | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      17879 TTTTGTGTTTTTTTATTTGTTTTTTGTTGTTGATTAGGAGATTAATAGT 17932

RESULT 12
US-10-424-599-102083/c
; Sequence 102083, Application US/10424599
; Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
;
; SEQ ID NO 102083

```

```

RESULT 12
US-10-424-599-102083/c
Sequence 102083, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yinhua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424.599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 102083
LENGTH: 1214
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1214)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_63198C.1
US-10-424-599-102083

Query Match
9.1%; Score 63.4; DB 12; Length 1214;
Best Local Similarity 46.9%; Pred. No. 6.5e-05;
Matches 157; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 325 TTCTGTCGATTCCTCTTTCTTCGCGTGAAAATGCGCCATGATTCGATTCG 384
Db 1169 TTTTTTNNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1110
QY 385 AAGTTTGTGCGATAGGTACTTTTCCATATTTATAGTCTTAGCAAGATA 444
Db 1109 NTMTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNNNTTT 1050
QY 445 CCGCGCTTACGTTTTTGTCAATTTGTGTCCTTACCGGTTAGTCGATCGA 504
Db 1049 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 990
QY 505 GATTGACGTGAAAATCCTTCGTTTTTGGTTTTTGTTCATATAAATCGATTGAT 564
Db 910 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTA 910

```

[illegible]

APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 37-21(10297)C
CURRENT APPLICATION NUMBER: US/09/983,965
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: US 09/465,231
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: US 60/113,678
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 5912
SEQ ID NO 2109
LENGTH: 529
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: unsure
LOCATION: (335)
OTHER INFORMATION:
OTHER INFORMATION: C]one ID: 34-LIB3057-015-Q1-K1-A6
US-09-983-965-2109

Query Match 9.0%; Score 62.8; DB 9; Length 529;
Best Local Similarity 49.0%; Pred. No. 6e-05;
Matches 166; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 337 TTTCTCTTTTCTCGCTGGAATAATGCGCTATGTCGATTCGAAGTTTGTG 396
DB |||||
53 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 112
QY 397 CTATGGGTACTTTTCCCTATTTTATGTTCTTAGTAAGATACGCGCTTAC 456
DB |||||
113 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 172
QY 457 TGTCTTGTCTATTTGTGCTTTCACCGTTTACGCGTATGCGAGTATGACTGT 516
DB |||||
173 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 232
QY 517 GAAATACCTGCTTTTGTGTTTGTCTATATAACGATGATGACCTTTGTG 576
DB |||||
233 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 292
QY 577 CTATGATGTTTGTGTTTGTAGCCTATGCGTTGTGCTGTATTAACCTGATG 636
DB |||||
293 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 352
QY 637 GTGATTTGAGATTTGTGAGTGTGAGTGTGCTTCTT 675
DB |||||
353 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 391

RESULT 14
US-10-125-968-78/c
Sequence 78, Application US/10125968
Publication No. US20030215805A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Palermo, Adam
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
APPLICANT: Elias, Josh
APPLICANT: Mertens, Maureen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-032
CURRENT APPLICATION NUMBER: US/10/125,968
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/285,163
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 1417
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 78

LENGTH: 629
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 25, 63, 64, 65, 142, 158, 159, 160, 204, 223, 224, 233, 255,
LOCATION: 256, 257, 258, 260, 263, 270, 271, 272, 286, 287, 290, 291,
LOCATION: 292, 293, 295, 296, 297, 298, 299, 303, 324, 333, 349,
LOCATION: 351, 352, 353, 354, 355, 356, 357, 358, 359, 360
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc.feature
LOCATION: 361, 362, 363, 364, 365, 366, 367, 369, 370, 371, 372,*373,
LOCATION: 374, 375, 376, 377, 380, 384, 386, 389, 390, 391, 393, 406,
LOCATION: 408, 409, 417, 419, 421, 424, 429, 434, 438, 439, 451, 453,
LOCATION: 456, 474, 475, 484, 486, 489, 492, 493, 501, 505, 516
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc.feature
LOCATION: 523, 536, 537, 538, 543, 546, 547, 554, 555, 558, 563, 565,
LOCATION: 568, 571, 572, 574, 575, 577, 581, 582, 586, 587, 592, 594,
LOCATION: 610, 611, 629
OTHER INFORMATION: n = A,T,C or G
US-10-125-968-78

Query Match 9.0%; Score 62.8; DB 15; Length 629;
Best Local Similarity 39.7%; Pred. No. 6.5e-05;
Matches 121; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 371 TGTCTCGATTTGGAAGTTTGTGCTATGCGTACTTTTCCCTATATTTATAGT 430
DB |||||
617 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 558
QY 431 CTGAGTAAAGTACTGCGCTTACTGTTTGTTCATTTGTGTGCTTACCGTT 490
DB |||||
557 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 498
QY 491 AGTCTGATCGAGATTTGACTGGAATAACCTGTTTGTGTTTGTTCAT 550
DB |||||
497 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 438
QY 551 TAAATGATGATCTACTTTTGTGCTTGTGATGTTGTTTGTAGCCTATGCTGT 610
DB |||||
437 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 378
QY 611 GCGTGTATTAATCTTACGTTGATGTCGATTTGAGATTTGTGAGTGTGCTT 670
DB |||||
377 NNN 318
QY 671 TCTTT 675
DB |||||
317 TTTT 313

RESULT 15
US-10-221-714A-502
Sequence 502, Application US/10221714A
Publication No. US20040048254A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with
TITLE OF INVENTION: tumor suppressor genes and oncogenes
FILE REFERENCE: 5013.1005
CURRENT APPLICATION NUMBER: US/10/221,714A
CURRENT FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: PCT/EP01/02955
PRIOR FILING DATE: 2001-03-15
PRIOR APPLICATION NUMBER: DE 10013847.0
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 14:30:53 ; Search time 4901.81 Seconds

(without alignments)
16605.743 Million cell updates/sec

Title: US-10-030-829-2

Perfect score: 1878
Sequence: 1 atgagcttcagagctgcgtcc.....ttcacatgagatgatga 1878

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GeneBml:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1878	100.0	1878	6 AX078761	AX078761 Sequence
2	1878	100.0	1878	6 BT004380	BT004380 Arabidops
3	1878	100.0	2162	8 BT002944	BT002944 Arabidops
4	1462	77.8	2254	8 AF239719	AF239719 Arabidops
5	1462	77.8	3275	6 AF078760	AF078760 Sequence
6	1462	77.8	8165	8 AB025633	AB025633 Arabidops
7	437.8	23.3	650	8 AT528171	AT528171 Arabidops
8	427.4	22.8	2181	8 AK064995	AK064995 Oryza sat
9	372	19.8	2570	8 AK100699	AK100699 Oryza sat
10	368	19.6	1556	8 AK064217	AK064217 Oryza sat
11	291.6	15.5	1564	8 AF542974	AF542974 Trilicium
12	288.8	15.4	1214	8 AK061975	AK061975 Oryza sat
13	192	10.2	513	8 CNG508CAO	AF469493 Trilicium
14	169.4	9.0	91053	8 CNG508CAO	AL831811 Oryza sat
15	169.4	9.0	112231	8 CNG507Y01	AL713901 Oryza sat
16	130.4	6.9	94911	8 AT114E10	AL18656 Arabidops
17	88.4	4.7	179685	10 AC122056	AC122056 Mus muscu
18	87.8	4.6	204635	10 AL928594	166494 Sequence 14
19	86.6	4.5	282861	2 AC111701	AL928594 Mouse DNA
20	84	4.4	227272	2 AC103478	AC111701 Rattus no
21	83.4	4.4	227272	2 AC103478	AC103478 Rattus no
22	83.2	4.4	279587	2 AC112408	AC112408 Rattus no
23	82.4	4.4	157393	10 AC124199	AC124199 Mus muscu
24	80.2	4.3	248672	2 AC105804	AC105804 Rattus no
25	79.4	4.2	124244	10 AL929026	AL929026 Mouse DNA
26	79.4	4.2	188782	10 AL713870	AL713870 Mouse DNA
27	79	4.2	195630	10 AC098726	AC098726 Mus muscu
28	78.6	4.2	114848	9 AC018664	AC018664 Homo sapi
29	78.6	4.2	178635	9 AC008080	AC008080 Homo sapi
30	78.6	4.2	194006	10 AL732441	AL732441 Mouse DNA
31	78	4.2	233713	2 AC133034	AC133034 Rattus no
32	78	4.2	237549	2 AC112451	AC112451 Rattus no
33	77.8	4.1	201074	2 AC102097	AC102097 Mus muscu
34	77.8	4.1	211947	10 AC107828	AC107828 Mus muscu
35	77.6	4.1	188587	10 AC124464	AC124464 Mus muscu
36	77.6	4.1	205037	10 AC126266	AC126266 Mus muscu
37	77.4	4.1	169256	10 AC127264	AC127264 Mus muscu
38	77.4	4.1	178927	2 AC118897	AC118897 Rattus no
39	77.4	4.1	217511	10 AC125357	AC125357 Mus muscu
40	77.2	4.1	56052	8 AP006363	AP006363 Lotus cor
41	77	4.1	215910	10 AC127335	AC127335 Mus muscu
42	76.8	4.1	195987	10 AC122537	AC122537 Mus muscu
43	76.8	4.1	197599	2 AC119769	AC119769 Rattus no
44	76.8	4.1	226601	10 AC122863	AC122863 Mus muscu
45	76.6	4.1	248718	10 AC117629	AC117629 Mus muscu

ALIGNMENTS

RESULT 1
AX078761 1878 bp DNA linear PAT 22-FEB-2001
LOCUS AX078761
DEFINITION Sequence 2 from Patent WO0105951.
ACCESSION AX078761
VERSION AX078761.1 GI:13158380
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
AUTHORS Beclin,C., Elmayer,T. and Vaucheret,H.
TITLE Novel gsg3 plant gene and use thereof

Pred. No. is the number of results predicted by chance to have a

JOURNAL

Patent: WO 0105951-A 2 25-JAN-2001;
 AVENTIS CROPS SCIENCE S.A.. (FR) ; INSTITUT NATIONAL DE LA RECHERCHE
 AGRONOMIQUE (FR)

FEATURES

Location/Qualifiers
 1..1878

/organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

CDS

1..1878
 /code="unnamed protein product"

/codon_start=1
 /protein_id="CAC32419.1"
 /db_xref="GI:13158381"

/translation="MSSRAGPMSKEKNVGGYRPEVEDLVGLAGTRLASQDGGEW
 EYISKNNKPGNTSGKTWVSONSPKRAMGQOQGRGNSVGRNNVGRNGRG
 IOANTSGRRLSRKYDNFVA PYPVSRPLEGMMWMOARGSAOHTAVQEPVEDD
 VNASREEDSDALDSDDDDLASDYPDSDVSOKSHGRKQNMKPKFSLSLIEO
 INEPOKWHCPACQNGPGLDWNHPLAHARTKARVYKRLHRELAEBLELDLMDRG
 ASVIPCGETYGGQKGLGDEKDEIVMPWVIMNTRLDKDNKMLGNOBLRYF
 DKREALRARSYGPQGRMSVLMESATGYLEARLHRELAEMGLDRIAWQKRM
 FSGVROLYGLATKODLDIPNHSQKRLFKELSYQEMVVKELRQISENDOOLNY
 FNKLSKONKHAUVLESELSKRLRTAEDNRIYRORTKQHEQNEEMDAHDFE
 MSIKOIHERRDAKENFEMLOQERAKVYVQOQOQINPSNDQCKRAEVSFIEF
 QEKEMFEVEREMLIKDEKMEKMKHHEEIPLEKEPDEALQLMYRKLHNEED
 D"

ORIGIN

Query Match 100.0%; Score 1878; DB 6; Length 1878;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1878; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAGTTCTAGGCTGCTCCAAATGCTTAAGAAAAGAAAGCTTCAGGCTGTTATAGCCT 60
DB 1 ATGAGTTCTAGGCTGCTCCAAATGCTTAAGAAAAGAAAGCTTCAGGCTGTTATAGCCT 60
QY 61 GAGGTTGAACAGTTGGTTCAAGTTGGCAGGAGCAGAGTGGCTTTCACAAGATGAT 120
DB 61 GAGGTTGAACAGTTGGTTCAAGTTGGCAGGAGCAGAGTGGCTTTCACAAGATGAT 120
QY 121 GAGAGAGATGGAGGTCTATTTCCAGAGAACAGAGAACAGAACCACTTCTGGA 180
DB 121 GAGAGAGATGGAGGTCTATTTCCAGAGAACAGAGAACAGAACCACTTCTGGA 180
QY 181 AAAAATTGGGTTTCTCAAAATCTCTCTAGAGCTTGGGTTGTCAGCAAGG 240
DB 181 AAAAATTGGGTTTCTCAAAATCTCTCTAGAGCTTGGGTTGTCAGCAAGG 240
QY 241 AGAGGTAGCAACGATCTGGGAGAGAAACAATGATCCGGAGAGGTAACGCAATGCT 300
DB 241 AGAGGTAGCAACGATCTGGGAGAGAAACAATGATCCGGAGAGGTAACGCAATGCT 300
QY 301 CGGGGCAATTCAGCTAAATATCTGTCGGGAGCAGAGCTTGAAGCAAAAGTATGATAC 360
DB 301 CGGGGCAATTCAGCTAAATATCTGTCGGGAGCAGAGCTTGAAGCAAAAGTATGATAC 360
QY 361 AACTTTGTGCAACCCCACTGTATCTGCGCCCTCTTGAAGAGAGATGGAATTGGCAG 420
DB 361 AACTTTGTGCAACCCCACTGTATCTGCGCCCTCTTGAAGAGAGATGGAATTGGCAG 420
QY 421 GCAAGAGAGGTTCTGCTCAGCACAACAGCTGTCAGAGAGTTCTCTGAGTGAAGATGAT 480
DB 421 GCAAGAGAGGTTCTGCTCAGCACAACAGCTGTCAGAGAGTTCTCTGAGTGAAGATGAT 480
QY 481 GTGATATATGCTTCTGAGAGAGAGATGATCCGATGCTTGAATGATTTGATGACGAC 540
DB 481 GTGATATATGCTTCTGAGAGAGAGATGATCCGATGCTTGAATGATTTGATGACGAC 540
QY 541 CTTCGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 CTTCGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 AATAAGTGCTTCAAAAAGTTCTTGGCAGCTTGGATGATGCTTGTGATGACAGATTAAT 660

```

```

DB 601 AATAAGTGCTTCAAAAAGTTCTTGGCAGCTTGGATGATGCTTGTGATGACAGATTAAT 660
QY 661 GAACCAAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 GAACCAAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 TATTAACCTGACCCCTCTACTAGCTCATGCGAGCAAAAAGAGATGACGAGTTAAGCTC 780
DB 721 TATTAACCTGACCCCTCTACTAGCTCATGCGAGCAAAAAGAGATGACGAGTTAAGCTC 780
QY 781 CATAGAGATTTGGCTGAAATTTTGAAGAAAGATCTACAGATGAGAGCGCATCTGTCAAT 840
DB 781 CATAGAGATTTGGCTGAAATTTTGAAGAAAGATCTACAGATGAGAGCGCATCTGTCAAT 840
QY 841 CTTGTGCTGAGATTTATGGCAGTGGAGGTTTGGGTGAGGATGAAAAGATTAAGAA 900
DB 841 CTTGTGCTGAGATTTATGGCAGTGGAGGTTTGGGTGAGGATGAAAAGATTAAGAA 900
QY 901 AATTGCTGGCTCCAAATGTCATCATGTAATCTAGACTGATTAAGAGATTAAGAT 960
DB 901 AATTGCTGGCTCCAAATGTCATCATGTAATCTAGACTGATTAAGAGATTAAGAT 960
QY 961 AAGTGGCTCGGATGAGGCAACCAAGAGCTGTAATCTTGAACAATGAGGCTCTT 1020
DB 961 AAGTGGCTCGGATGAGGCAACCAAGAGCTGTAATCTTGAACAATGAGGCTCTT 1020
QY 1021 AGAGCAGCCATTCCTATGCTCCACAGAGCCATGCTGGAGAGAGGTTCTGATGTTAG 1080
DB 1021 AGAGCAGCCATTCCTATGCTCCACAGAGCCATGCTGGAGAGAGGTTCTGATGTTAG 1080
QY 1081 AGCAGTGCACATGCTATTTGAGAGCCGACCTCCACCGGAGTTAGCTGAGTGGG 1140
DB 1081 AGCAGTGCACATGCTATTTGAGAGCCGACCTCCACCGGAGTTAGCTGAGTGGG 1140
QY 1141 TTAGATGAATTTGCTGGGCTGAGAGGCGAGATGTTTCTGAGAGTGTTCGCCAATG 1200
DB 1141 TTAGATGAATTTGCTGGGCTGAGAGGCGAGATGTTTCTGAGAGTGTTCGCCAATG 1200
QY 1201 TATGCTTCCTTGAACAGAGCAAGCAATCTGAGCATATTCATCAACATCTCAAGCAAA 1260
DB 1201 TATGCTTCCTTGAACAGAGCAAGCAATCTGAGCATATTCATCAACATCTCAAGCAAA 1260
QY 1261 ACAAGGCTGAATTTGAGTTGAATCATACAGAGATGTTTGAAGAGAGCTGAGCGAG 1320
DB 1261 ACAAGGCTGAATTTGAGTTGAATCATACAGAGATGTTTGAAGAGAGCTGAGCGAG 1320
QY 1321 ATCTCTGAGCAATCTGAGAGCTGAATCTTGAATTTATGAGAGAGCTGAGCGAG 1380
DB 1321 ATCTCTGAGCAATCTGAGAGCTGAATCTTGAATTTATGAGAGAGCTGAGCGAG 1380
QY 1381 CAGCGCAAGGCTGAGAGATCTCTGAAATTTATGAGAGAGAGCTGAGCGAG 1440
DB 1381 CAGCGCAAGGCTGAGAGATCTCTGAAATTTATGAGAGAGAGCTGAGCGAG 1440
QY 1441 GAGGATATCGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 1441 GAGGATATCGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 ATGATGACACAGCAGGTTTTCATGATTAATCAATCAATCAATCAATCAATCAATCA 1560
DB 1501 ATGATGACACAGCAGGTTTTCATGATTAATCAATCAATCAATCAATCAATCAATCA 1560
QY 1561 GCAAGAGAGAGATTTGAGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 1561 GCAAGAGAGAGATTTGAGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 CAGCAGAGAGAGATTTATCTTACAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1680
DB 1621 CAGCAGAGAGAGATTTATCTTACAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1680
QY 1681 TCAAGCTTATGAGATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740

```


Db 1681 TCAAGCTTCATCGAGTTTCAAGAGAAAAGAGATGAGAGAGTTGTGAGAGAGAGGAGATG 1740

Qy 1741 CTGATTAAGATCAAGAGAAAGATGAGATGAAAGAGGATCAGAGAGATG 1800

Db 1741 CTGATTAAGATCAAGAGAAAGATGAGATGAAAGAGGATCAGAGAGATG 1800

Qy 1801 TTGTATCTGGAGAGAAATTTGATGAGGCTTTGGAACAGCTCATGTACAAGATGGCCTT 1860

Db 1801 TTGTATCTGGAGAGAAATTTGATGAGGCTTTGGAACAGCTCATGTACAAGATGGCCTT 1860

Qy 1861 CACATGAAGATGATGCA 1878

Db 1861 CACATGAAGATGATGCA 1878

RESULT 2

BT004380 1909 bp mRNA linear PLN 14-FEB-2003

LOCUS Arabidopsis thaliana clone U20243 unknown protein (At5g23570) mRNA,

DEFINITION complete cds.

ACCESSION BT004380

VERSION BT004380.1 GI:28393932

KEYWORDS FLI_CDN.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustoids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1909)

AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinzaki,K., Davis,R.W., Becker,J.R. and Theologis,A.

TITLE Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1909)

AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinzaki,K., Davis,R.W., Becker,J.R. and Theologis,A.

TITLE Direct Submission

JOURNAL Submitted (14-FEB-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFLP CDNA (RFLP CDNA: RIKEN Arabidopsis Full-Length CDNA): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinzaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNT (ORF) clones using the RFLP CDNA: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinzaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

Location/Qualifiers

1..1909

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="U20243"

/ecotype="Columbia"

/note="this clone is in pUNT 51."

1..1909

/gene="At5g23570"

1..1878

/gene="At5g23570"

/codon_start=1

/evidence=experimental

/product="unknown protein"

/protein_id="AA042374.1"

/db_xref="GI:28393932"

/translation="MSSRAGPWSKERNVGGYRPPEVQLVGLAGTRLLASODDGESEVVISKNNKRPENTSGKTIVSGNSNPAPRWGQOQGRGNSVGRGNVSGRGNNGRGIOANISGRGRLSRKTDNNFVAPPPVSRPPLSGNMWQARGSAGHTAVQEPFVDEDVNABEENDSDALDDSDDLASDVIDSVSQSHSGSRKQNMFKFKFSLDLSLEQINEPOROMHPACONGPGLADWYNLPLAHARTGARVKLHRELAEVLERDLQWRASVIPCGETYGGWKGLGEDEKDYETIWPMPVITIMNTRLDKDNDKMLGNGBLEYFDEVALRAHRSYGPQGHGMSVLMFESSATGYLEAERLHRELAEMLDRILANGQKSMBSGTVRQLYGLATKODLDIPNHSQGRRLKFLKSYEMVKELRQISENQNLNFFNKLISKQNHAKVLEESLEIVSEKLRRTAEDNRIVRQTKQHEQNRSEMDHDEFMSIKQIHERDRPAKSENFEMLQOQERAKVVGQOQONINSSNDCKRKEVSSFLFQEKEMBEFVEREMLIKDQEKMEDMKKRHHEIFDLEREFDALAEQLMYKQLHNEDD"

1879..1909

/gene="At5g23570"

ORIGIN

3'UTR

Query Match 100.0%; Score 1878; DB 8; Length 1909;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1878; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTTCTAGAG3CTGTGTCCTCAATGTCCTAAGAAAAGACGTTCAAGGCTGTTATGAGCCT 60

Db 1 ATGAGTTCTAGAG3CTGTGTCCTCAATGTCCTAAGAAAAGACGTTCAAGGCTGTTATGAGCCT 60

Qy 61 GAGGTTGAACACTTGGTTCAGAGTTTGGAGGAGAGAGAGCTGGCTTCCACAAGTAT 120

Db 61 GAGGTTGAACACTTGGTTCAGAGTTTGGAGGAGAGAGAGCTGGCTTCCACAAGTAT 120

Qy 121 GGAGGAGAGTGGGAGAGCTATTTCCAAAGAAACAAGAACAGGAAACACTTCTGGA 180

Db 121 GGAGGAGAGTGGGAGAGCTATTTCCAAAGAAACAAGAACAGGAAACACTTCTGGA 180

Qy 181 AAAACTGGGTTTCTCAGAAATTCCTCTAGAGCTTGG3GTGTGACGACGAGG 240

Db 181 AAAACTGGGTTTCTCAGAAATTCCTCTAGAGCTTGG3GTGTGACGACGAGG 240

Qy 241 AAGAGTGAACAAGTATCTGGGAGAGAAACAATGTTATCCGGAGAGGTAAACGGCAATGCT 300

Db 241 AAGAGTGAACAAGTATCTGGGAGAGAAACAATGTTATCCGGAGAGGTAAACGGCAATGCT 300

Qy 301 CGGGGCAATCAAGCTAACATATCTGGTGGGAGACGAGCTTGAGCGAAGATGATAAC 360

Db 301 CGGGGCAATCAAGCTAACATATCTGGTGGGAGACGAGCTTGAGCGAAGATGATAAC 360

Qy 361 AACTTTGTGGACCCCACTGATCTCGCCCTCTTTGGAAGAGAGATGAAATTTGGCAG 420

Db 361 AACTTTGTGGACCCCACTGATCTCGCCCTCTTTGGAAGAGAGATGAAATTTGGCAG 420

Qy 421 GCAAGAGAGGTTTGTCTCAGACACAGTGTGACAGAGATTTCTCGAAGTGAAGATGAT 480

Db 421 GCAAGAGAGGTTTGTCTCAGACACAGTGTGACAGAGATTTCTCGAAGTGAAGATGAT 480

Qy 481 GTGATTAATGCTTCTGAGAGAGAGATGTTCCGATGCTTTGGATGATTTCTGATGACGAC 540

Db 481 GTGATTAATGCTTCTGAGAGAGAGATGTTCCGATGCTTTGGATGATTTCTGATGACGAC 540

Qy 541 CTTGCAAGTGAATGATTAATGATCTCGAGATGTGATCAAAAGCCATGATCAAGAAAGCAG 600

Db 541 CTTGCAAGTGAATGATTAATGATCTCGAGATGTGATCAAAAGCCATGATCAAGAAAGCAG 600

QY 601 AATAAGTGTTCAAAAGTCTTTGGGACGCTTGGAATGCTTGATGAGAGAGATTAAT 660
 DB 601 AATAAGTGTTCAAAAGTCTTTGGGACGCTTGGAATGCTTGATGAGAGAGATTAAT 660
 QY 661 GAACCAAGAGGAGGAGTGTTCAGAGTGTTCAGAAAGGAGCTGTGCTCATGATGG 720
 DB 661 GAACCAAGAGGAGGAGTGTTCAGAGTGTTCAGAAAGGAGCTGTGCTCATGATGG 720
 QY 721 TATTAACCTGACACCTCTTCTAAGCTCACTGAGAGCAAAAGAGCTTAGGAGTTAACTC 780
 DB 721 TATTAACCTGACACCTCTTCTAAGCTCACTGAGAGCAAAAGAGCTTAGGAGTTAACTC 780
 QY 781 CATTAAGAGATTTGGCTGAATTTTAAAGAGATCTAAGATGAGAGGCGCATCTGCAAT 840
 DB 781 CATTAAGAGATTTGGCTGAATTTTAAAGAGATCTAAGATGAGAGGCGCATCTGCAAT 840
 QY 841 CCTTGTGTGAGATTTATGAGGAGTGTGGAAGGCTTTGGTGAAGATGAAAGATTAATGA 900
 DB 841 CCTTGTGTGAGATTTATGAGGAGTGTGGAAGGCTTTGGTGAAGATGAAAGATTAATGA 900
 QY 901 ATTGTGTGCTTCGAATGCTCATCATGATTAATCTAAGCTGATTAAGAGCATTAACGAT 960
 DB 901 ATTGTGTGCTTCGAATGCTCATCATGATTAATCTAAGCTGATTAAGAGCATTAACGAT 960
 QY 961 AAGTGTGCTGAGATGAGGAGCAAGAGGCTGTGGAATCTTCAAGAGATTAAGAGCTCTT 1020
 DB 961 AAGTGTGCTGAGATGAGGAGCAAGAGGCTGTGGAATCTTCAAGAGATTAAGAGCTCTT 1020
 QY 1021 AGAGCAAGCATTCCTATGAGTCCAGAGGCGCATCTGAGATGAGTGTTCATGATTTAG 1080
 DB 1021 AGAGCAAGCATTCCTATGAGTCCAGAGGCGCATCTGAGATGAGTGTTCATGATTTAG 1080
 QY 1081 AGCAGTGTGCACTGAGTGTTCAGAGGCGCATCTGAGATGAGTGTTCATGATTTAG 1140
 DB 1081 AGCAGTGTGCACTGAGTGTTCAGAGGCGCATCTGAGATGAGTGTTCATGATTTAG 1140
 QY 1141 TTAGATTAAGATTTGCTGGGCTGAGAGGCGCATCTGAGATGAGTGTTCATGATTTAG 1200
 DB 1141 TTAGATTAAGATTTGCTGGGCTGAGAGGCGCATCTGAGATGAGTGTTCATGATTTAG 1200
 QY 1201 TATGCTTCTCTTCAAGCAAGAGATCTGAGATTAATCAATCAACACTCTCAAGGCAAA 1260
 DB 1201 TATGCTTCTCTTCAAGCAAGAGATCTGAGATTAATCAATCAACACTCTCAAGGCAAA 1260
 QY 1261 ACAAGGCTGAATTCGAGTGTTCAGAGGCGCATCTGAGATGAGTGTTCATGATTTAG 1320
 DB 1261 ACAAGGCTGAATTCGAGTGTTCAGAGGCGCATCTGAGATGAGTGTTCATGATTTAG 1320
 QY 1321 ATCTCTGAGAGCAATGAGAGCTGAGATTAATCAATCAACACTCTCAAGGCAAA 1380
 DB 1321 ATCTCTGAGAGCAATGAGAGCTGAGATTAATCAATCAACACTCTCAAGGCAAA 1380
 QY 1381 CACGCCAGGCTGTGAGATTCCTGGAATTAATGAGAGAGGCTGAGAGGCAAG 1440
 DB 1381 CACGCCAGGCTGTGAGATTCCTGGAATTAATGAGAGAGGCTGAGAGGCAAG 1440
 QY 1441 GAGATTAATGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 DB 1441 GAGATTAATGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 QY 1501 ATGAGATCAACAGAGAGGCTTTTCAATGATTAATCAATCAACAGATTCATGAAGAGAG 1560
 DB 1501 ATGAGATCAACAGAGAGGCTTTTCAATGATTAATCAATCAACAGATTCATGAAGAGAG 1560
 QY 1561 GCAAGAGAGAGAGATTTTCAATGATTTTCAATGATTTTCAATGATTTTCAATGATTT 1620
 DB 1561 GCAAGAGAGAGAGATTTTCAATGATTTTCAATGATTTTCAATGATTTTCAATGATTT 1620
 QY 1621 CAGAGAGAGAGAGATTTTCAATGATTTTCAATGATTTTCAATGATTTTCAATGATTT 1680
 DB 1621 CAGAGAGAGAGAGATTTTCAATGATTTTCAATGATTTTCAATGATTTTCAATGATTT 1680

QY 1681 TCAAGCTTCATGAGATTTCAAGAGAGAGAGAGAGAGATTTTGTGAAGAGAGAGATG 1740
 DB 1681 TCAAGCTTCATGAGATTTCAAGAGAGAGAGAGAGAGATTTTGTGAAGAGAGAGATG 1740
 QY 1741 CTGATTAAGATTCAGAGAGAGAGAGAGAGATTCATGAAGAGAGAGAGAGAGATTA 1800
 DB 1741 CTGATTAAGATTCAGAGAGAGAGAGAGAGATTCATGAAGAGAGAGAGAGAGATTA 1800
 QY 1801 TTTGATCTGAGAGAGAGATTTGATGAGGCTTTGGAACAGCTCATGTCAACATGAGCTT 1860
 DB 1801 TTTGATCTGAGAGAGAGATTTGATGAGGCTTTGGAACAGCTCATGTCAACATGAGCTT 1860
 QY 1861 CACATGAAGATGATGA 1878
 DB 1861 CACATGAAGATGATGA 1878
 RESULT 3
 BT002944 2162 bp mRNA linear PLN 15-JAN-2003
 LOCUS Arabidopsis thaliana clone RAFL14-93-K05 (R20243) unknown protein
 DEFINITION (AT5G23570) mRNA, complete cds.
 ACCESSION BT002944 GI:27754622
 VERSION BT002944.1
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M.,
 Onodera, C.S., Quach, H.L., Tang, C.C., Tortum, M., Wong, C., Wu, H.C.,
 Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,
 Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P.,
 Southwick, A., Tripp, M.G., Wu, T., Shinzaki, K., Davis, R.W.,
 Ecker, J.R. and Theologis, A.
 TITLE Arabidopsis Full Length cDNA Clones
 JOURNAL Arabidopsis Full Length cDNA Clones
 REFERENCE
 AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M.,
 Onodera, C.S., Quach, H.L., Tang, C.C., Tortum, M., Wong, C., Wu, H.C.,
 Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,
 Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P.,
 Southwick, A., Tripp, M.G., Wu, T., Shinzaki, K., Davis, R.W.,
 Ecker, J.R. and Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
 Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinzaki, K.
 The Salk, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M.,
 Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M., Onodera, C.S.,
 Quach, H.L., Tang, C.C., Tortum, M., Wong, C., Wu, H.C., Yu, G.,
 Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,
 Palm, C.J., Shim, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,
 Ecker, J.R. and Theologis, A.
 Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)
 contributed equally to this work as PIs.
 Annotation based on July 2002 version of the Arabidopsis genome
 submitted to Genbank.
 FEATURES
 Location/Qualifiers
 1..2162
 source

```
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAF1.4-93-K05 (R20243)"
/ecotype="Columbia"
/notes="This clone is in a modified pBluescript vector2
(lambda PS) as a BamHI/XhoI insert."
1..2162
/gene="At5g23570"
1..91
/gene="At5g23570"
92..1969
/gene="At5g23570"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA02757.1"
/db_xref="GI:2754623"
/tranlation="MSSRAGPMKRNKVGQYRPEVQLVGLAGTRIASDDGEM
EVISSKRNKRGSTGKTVWSNPNPRAAGQOQGRGSVSGRNNVSGRGNGRG
IQANISGRGRLSRKTDNPFVAPRPPVSRPLBEGMNMQARSGAHTVADEFVEDD
VDVASSEENDSDALDSDDDLDSDYDSDVSQSHSRKQKPFPSGLDLSLEQ
INERPOQWCPACONGPGLIDWYNLHLPLAHARTKGRVYKLELAEVLQMRG
ASVIPCSEIYGQKGLGEDEKQYEIVWPVWVIMNTRLDKODNKLGMGNOLLEYF
DKREALRARSVGRPOGRHGMVLPSSATGVLEARIHRELAEMLDRIAMGOKRSM
FSGVRLQYGLFLTKODDIFNHSOGKTRKLELSQEMVYKELRQISENQOLNY
FKNKLKONKHAVLESLEINSEKLRTRADNRIVQRTRQWHEQREBMDHDFE
MOSIKQIHERDAKENFEMLOQERAKVVGQOQNIINPSNDCKRAEVSFTF
OEKEMEFPVEEREMLIKDEKMEWMKRRHEEIPLEKEFDEALBOLMYKHGLNED
D"
1970..2162
/gene="At5g23570"
2166
/gene="At5g23570"
/notes="not present in genomic sequence"

ORIGIN
Query Match 100.0%; Score 1878; DB 8; Length 2162;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTTCTAGGCTGGTCCAAATGCTAAAGAAAAGAACCTTCAGGCTGTATAGGCT 60
Db 92 ATGAGTTCTAGGCTGGTCCAAATGCTAAAGAAAAGAACCTTCAGGCTGTATAGGCT 151
QY 61 GAGGTGAACAGTTGTTCAAGGTTGGCAGGAGCGAGACTGCTCTTCAAGATGAT 120
Db 152 GAGGTGAACAGTTGTTCAAGGTTGGCAGGAGCGAGACTGCTCTTCAAGATGAT 211
QY 121 GAGGAGAGTGGAGGTCATTTCCAGAGAAACAAGAACCAAGAACCATTTCTGGA 180
Db 212 GAGGAGAGTGGAGGTCATTTCCAGAGAAACAAGAACCAAGAACCATTTCTGGA 271
QY 181 AAAAATTGGGTTCTCAGAAATTCCTCTAGAGCTTGGGTTGTCAGCAGCAAGG 240
Db 272 AAAAATTGGGTTCTCAGAAATTCCTCTAGAGCTTGGGTTGTCAGCAGCAAGG 331
QY 241 AGAGGTAGCAACGTATCTGGAGAGGAAACAATGTATCCGAGAGAGTTAACCGCAATG 300
Db 332 AGAGGTAGCAACGTATCTGGAGAGGAAACAATGTATCCGAGAGAGTTAACCGCAATG 391
QY 301 CGGGGCAATTAACCTAATATCTGGTCGGGGAGCAGGCTTGAGCAGAAAGTATGATTA 360
Db 392 CGGGGCAATTAACCTAATATCTGGTCGGGGAGCAGGCTTGAGCAGAAAGTATGATTA 451
QY 361 AACTTTGGCACCACCACTGTATCTCGCCCTCTTTGGAAGAGGATGAATTTGGCAG 420
Db 452 AACTTTGGCACCACCACTGTATCTCGCCCTCTTTGGAAGAGGATGAATTTGGCAG 511
QY 421 GCAAGAGAGGTTCTGTCAGACACAGCTGTGCAAGAGATTCTGACGTTGAGAGATG 480
Db 512 GCAAGAGAGGTTCTGTCAGACACAGCTGTGCAAGAGATTCTGACGTTGAGAGATG 571
```

```
QY 481 GTGATTAATGCTTCTGAGGAAGAAATGATTCGATGCTTTGGATGATTTCTGATGACGAC 540
Db 572 GTGATTAATGCTTCTGAGGAAGAAATGATTCGATGCTTTGGATGATTTCTGATGACGAC 631
QY 541 CTTGCAAGTATGATTAATGATCTCGAGTGTGAGTCAAAAGACCATGATCAAGAAACGAG 600
Db 632 CTTGCAAGTATGATTAATGATCTCGAGTGTGAGTCAAAAGACCATGATCAAGAAACGAG 691
QY 601 AATAAGTGTGCAAAAAGTCTTTGGCAGCTTGATGATGATGATGATGATGATGATGAT 660
Db 692 AATAAGTGTGCAAAAAGTCTTTGGCAGCTTGATGATGATGATGATGATGATGATGAT 751
QY 661 GAACCAAGAGGACGATGATGATTCAGGCTTGACAGAAAGGACCTGTCATGATGATGAT 720
Db 752 GAACCAAGAGGACGATGATGATTCAGGCTTGACAGAAAGGACCTGTCATGATGATGAT 811
QY 721 TATTAACCTGACCCCTTACTAGCTCATGCGAGAGCAAAAAGAGCTTAGGCGATTAAGCTC 780
Db 812 TATTAACCTGACCCCTTACTAGCTCATGCGAGAGCAAAAAGAGCTTAGGCGATTAAGCTC 871
QY 781 CATAGGAATTTGCTGAAGTTTAAAGAAAGATCTACAGATGAGAGCCGATCTGTCAAT 840
Db 872 CATAGGAATTTGCTGAAGTTTAAAGAAAGATCTACAGATGAGAGCCGATCTGTCAAT 931
QY 841 CCTTGGTGAAGATTTATGGGCGATGGAAGGGTTTGGGTGAGGATGAAAAGATTATGAA 900
Db 932 CCTTGGTGAAGATTTATGGGCGATGGAAGGGTTTGGGTGAGGATGAAAAGATTATGAA 991
QY 901 ATTGTCGCTCCCAATGATGATCATCATGAATTAATGATGATGATGATGATGATGAT 960
Db 992 ATTGTCGCTCCCAATGATGATCATCATGAATTAATGATGATGATGATGATGATGAT 1051
QY 961 AATGCTGCTGGCATGAGGCAACCAAGGCTGCTGGAATCTTGCAAGATATGAGGCTCTT 1020
Db 1052 AATGCTGCTGGCATGAGGCAACCAAGGCTGCTGGAATCTTGCAAGATATGAGGCTCTT 1111
QY 1021 AGAGCAGGCAATTCATGATGATCCAGAGGCAATCGTGGAGTGAAGTGTTCATGATTTGAG 1080
Db 1112 AGAGCAGGCAATTCATGATGATCCAGAGGCAATCGTGGAGTGAAGTGTTCATGATTTGAG 1171
QY 1081 AGCAGTGCACATGCTGATTTTGAAGGCGCAAGCGCTCACCGGAGTTAGCTGATGAGTGG 1140
Db 1172 AGCAGTGCACATGCTGATTTTGAAGGCGCAAGCGCTCACCGGAGTTAGCTGATGAGTGG 1231
QY 1141 TTGATAGAAATTTGCTGGGCTCAGAAAGCGCAAGTATGTTTCTGAGAGTGTTCGCAACTG 1200
Db 1232 TTGATAGAAATTTGCTGGGCTCAGAAAGCGCAAGTATGTTTCTGAGAGTGTTCGCAACTG 1291
QY 1201 TATGCTTCTCTGCAAGCAAGCAAGATCTGGAATCTGAATCAATCAACTCTCAAGGCAAA 1260
Db 1292 TATGCTTCTCTGCAAGCAAGCAAGATCTGGAATCTGAATCAATCAACTCTCAAGGCAAA 1351
QY 1261 ACAAGGCTGAATTTGAGTTGAATATCATCAAGAGATGTTGTAAAGAGCTGAGGCAAG 1320
Db 1352 ACAAGGCTGAATTTGAGTTGAATATCATCAAGAGATGTTGTAAAGAGCTGAGGCAAG 1411
QY 1321 ATCTTGAGGACATGAGAGCTGAATCTTAAGAACAGAGCTTCAAAAACGAAACAG 1380
Db 1412 ATCTTGAGGACATGAGAGCTGAATCTTAAGAACAGAGCTTCAAAAACGAAACAG 1471
QY 1381 CACGCCAAGTGTGTTAGGAATCTCTGGAATTTATAGGAGAGAGGCTGGTGAATCTGCA 1440
Db 1472 CACGCCAAGTGTGTTAGGAATCTCTGGAATTTATAGGAGAGAGGCTGGTGAATCTGCA 1531
QY 1441 GAGGATTAATCGATCTGAGACAGAGAACTAAGATGACAGATGAACAGAACAGGAAGAG 1500
Db 1532 GAGGATTAATCGATCTGAGACAGAGAACTAAGATGACAGATGAACAGAACAGGAAGAG 1591
QY 1501 ATGATGACACAGACAGGTTTTTTCATGATTCATTAACAGATCCATGAAGAGAGAGAC 1560
Db 1592 ATGATGACACAGACAGGTTTTTTCATGATTCATTAACAGATCCATGAAGAGAGAGAC 1651
```

QY 1561 GCMAAGAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1620
 DB 1652 GCMAAGAGAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1711
 QY 1621 CAGCAGCAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1680
 DB 1712 CAGCAGCAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1771
 QY 1681 TCAGGCTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1740
 DB 1772 TCAGGCTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1831
 QY 1741 CTGATTAAGATCAGAGAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1800
 DB 1832 CTGATTAAGATCAGAGAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1891
 QY 1801 TTGATCTGAGAGAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1860
 DB 1892 TTGATCTGAGAGAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1951
 QY 1861 CACATGAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1878
 DB 1952 CACATGAGAGATTTCAGATGTTTGCAGCAGCAGGAACGCGCAAGGTTGTGCGCAG 1969

RESULT 4
 AF239719 2254 bp DNA linear PLN 02-JUN-2000
 LOCUS Arabidopsis thaliana SGS3 gene, complete cds.
 DEFINITION Arabidopsis thaliana SGS3 gene, complete cds.
 ACCESSION AF239719
 VERSION AF239719.1 GI:8164029
 KEYWORDS

SOURCE Arabidopsis thaliana (chale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.
 1 (bases 1 to 2254)
 Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C.,
 Morel, J.B., Jonette, D., Jacombe, A.M., Nikic, S., Picault, N.,
 Remoue, K., Sanial, M., Vo, T.A. and Vaucheret, H.
 Arabidopsis SGS2 and SGS3 genes are required for
 posttranscriptional gene silencing and natural virus resistance
 Cell 101 (5), 533-542 (2000)

TITLE
 JOURNAL MEDLINE 20306668
 PUBMED 10850495
 REFERENCE 2 (bases 1 to 2254)
 AUTHORS Beclin, C., Mourrain, P., Vaucheret, H. and Elmayan, T.
 TITLE Direct Submission
 JOURNAL Submitted (28-FEB-2000) Biologie Cellulaire, INRA, Route de
 Saint-Cyr, Versailles 78026, France
 Location/Qualifiers

FEATURES
 source
 1..2254
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultiivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="V"
 join(1..963,1037..1328,1440..1684,1787..1953,2044..2254)
 /product="SGS3"
 join(1..963,1037..1328,1440..1684,1787..1953,2044..2254)
 /codon_start=1
 /product="SGS3"
 /protein_id="AA073960.1"
 /db_xref="GI:8164030"
 /translation="MSSRAGPMSEKKNVQGYRPEVQLVQIAGRLASODDGEW
 EVISKRNKRGNTSGKTWVSQNSNPRAGQOQSGVSRGNVSGRNGRG
 IOANISGRGALSRKYDNNEVAPPPVSRPELEGWMAQAGSAGHTAVOEPDUD
 VDVAEENSDALDDSDDDLASPDYSDVSQSHSRKQKWFPSLSLSLEO
 INEPORMHCACONGGALDWYLIHPLHARTKGRVYKLRHLEAVLEKDLQMRG
 ASVPCGEITGVGGLGDEKDYIIVMPVYIMNTTLDKDNKMLGNGEELFV
 DYEALRRHSYTGQGRHGRMSVIMFESSATGYLEERHLEMLGIDRIAGOKRSM
 FSGVRDLVGLATKDDLDIFNHSQKTKLKLKLSYQGMVVKELRQISENQCLNY

ORIGIN

Query Match 77.8%; Score 1462; DB 8; Length 2254;
 Best Local Similarity 83.3%; Pred. No. 0;
 Matches 1878; Conservative 0; Mismatches 0; Indels 376; Gaps 4;

1 ATGAGTTTGAAGGCTGCTGCAATGCTTAAAGAAAGAAAGGTTCAAGGTTGTTAAGGCT 60
 1 ATGAGTTTGAAGGCTGCTGCAATGCTTAAAGAAAGAAAGGTTCAAGGTTGTTAAGGCT 60
 61 GAGTTGAACAGTTGGTTCAAGTTGGCAGGACGAGCTGCTTTTCAAGATGAT 120
 61 GAGTTGAACAGTTGGTTCAAGTTGGCAGGACGAGCTGCTTTTCAAGATGAT 120
 61 GAGTTGAACAGTTGGTTCAAGTTGGCAGGACGAGCTGCTTTTCAAGATGAT 120
 121 GAGGAGAGTGGAGAGTCAATTTCAAGAAAGAACAGAAACAAACCGAGAAACCTTCTGA 180
 121 GAGGAGAGTGGAGAGTCAATTTCAAGAAAGAACAGAAACAAACCGAGAAACCTTCTGA 180
 121 GAGGAGAGTGGAGAGTCAATTTCAAGAAAGAACAGAAACAAACCGAGAAACCTTCTGA 180
 181 AAAAATTGGTTTCTGCAATTCCTTCTGAGCTTGGGTTGTCAGACGAGG 240
 181 AAAAATTGGTTTCTGCAATTCCTTCTGAGCTTGGGTTGTCAGACGAGG 240
 181 AAAAATTGGTTTCTGCAATTCCTTCTGAGCTTGGGTTGTCAGACGAGG 240
 241 AGAGTGAACAGTATCTGGAGAGAGAAACATGATTCGGGAGAGTAAACGGCAATGAT 300
 241 AGAGTGAACAGTATCTGGAGAGAGAAACATGATTCGGGAGAGTAAACGGCAATGAT 300
 301 CGGGGCAATTCAGCTAACATTCCTGCTGGGAGAGAGCTTGAAGAGATGATTAAC 360
 301 CGGGGCAATTCAGCTAACATTCCTGCTGGGAGAGAGCTTGAAGAGATGATTAAC 360
 301 CGGGGCAATTCAGCTAACATTCCTGCTGGGAGAGAGCTTGAAGAGATGATTAAC 360
 361 AACTTTGGGCAACCCCACTGATCTCGCCCTCTTTGAAGAGATGATTAAC 420
 361 AACTTTGGGCAACCCCACTGATCTCGCCCTCTTTGAAGAGATGATTAAC 420
 361 AACTTTGGGCAACCCCACTGATCTCGCCCTCTTTGAAGAGATGATTAAC 420
 421 GCAAGAGAGAGTTTCTGCTCAGACACAGCTGTGAGAGAGTTTCTGAGAGAGATGAT 480
 421 GCAAGAGAGAGTTTCTGCTCAGACACAGCTGTGAGAGAGTTTCTGAGAGAGATGAT 480
 421 GCAAGAGAGAGTTTCTGCTCAGACACAGCTGTGAGAGAGTTTCTGAGAGAGATGAT 480
 481 GGGAGTAATGCTTCTGAGAGAGAGATGATTCGATGTTGATGATTTCTGATGAGAC 540
 481 GGGAGTAATGCTTCTGAGAGAGAGATGATTCGATGTTGATGATTTCTGATGAGAC 540
 481 GGGAGTAATGCTTCTGAGAGAGAGATGATTCGATGTTGATGATTTCTGATGAGAC 540
 541 CTTGCAAGTATGATTAATGATCTCGAGTGTGATGATTAAGAGAGATGATTAAGAG 600
 541 CTTGCAAGTATGATTAATGATCTCGAGTGTGATGATTAAGAGAGATGATTAAGAG 600
 541 CTTGCAAGTATGATTAATGATCTCGAGTGTGATGATTAAGAGAGATGATTAAGAG 600
 601 AATTAAGTGTGTAAGAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGAT 660
 601 AATTAAGTGTGTAAGAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGAT 660
 601 AATTAAGTGTGTAAGAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGAT 660
 661 GAACACAGAGAGAGTGTGCTGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 661 GAACACAGAGAGAGTGTGCTGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 661 GAACACAGAGAGAGTGTGCTGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 721 TATTAAGTGTGTAAGAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGAT 780
 721 TATTAAGTGTGTAAGAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGAT 780
 721 TATTAAGTGTGTAAGAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGAT 780
 781 CATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 781 CATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 841 CTTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 841 CTTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 841 CTTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 901 ATTGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 901 ATTGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960

```

QY 961 A----- 961
Db 961 AAGTGAATCTCTCTCTTACTTCTTAATTTTCTCTGCAATTCAGATCTTA 1020
QY 962 -----AGTGGCTGGCATGGGCAACCAAGAGCTGTGAAATCTTGCACAA 1007
Db 1021 GAATGTACATTTGATGGCTCGGCATGGGCAACCAAGAGCTGTGAAATCTTGCACAA 1080
QY 1008 GTATGAGGCTCTAGAGCAGCCATTCCTATGCTCCAGGGCCTGCTGGATGAGTGT 1067
Db 1081 GTATGAGGCTCTAGAGCAGCCATTCCTATGCTCCAGGGCCTGCTGGATGAGTGT 1140
QY 1068 TCTGATGTTGAGAGCAGTGCCTGCTATTTGAGGCGGAAAGCTCCACCGGAGTT 1127
Db 1141 TCTGATGTTGAGAGCAGTGCCTGCTATTTGAGGCGGAAAGCTCCACCGGAGTT 1200
QY 1128 AGCTGAGTGGGTTAGATGAAATTCCTGGGGTCAAGAGCGCATGTTTCTGGAGG 1187
Db 1201 AGCTGAGTGGGTTAGATGAAATTCCTGGGGTCAAGAGCGCATGTTTCTGGAGG 1260
QY 1188 TGTTCGCAACGTATGAGCTTCTTGCAGAGAGCAAGATCTGACATATTCATCAACA 1247
Db 1261 TGTTCGCAACGTATGAGCTTCTTGCAGAGAGCAAGATCTGACATATTCATCAACA 1320
QY 1248 CTCTCA----- 1253
Db 1321 CTCTCAAGTCTCTCCCAAAAGAAATTTGATATATGCTTTTGTGATGGAAT 1380
QY 1254 -----AGG 1256
Db 1381 TTAAGTTTGTGTCCGTGTTATGATCATCTGTATATATATCATGATTCATTAGG 1440
QY 1257 CAAAACAGGCTGAATTCAGATTTGAATCATACCAAGAGATGGTTGTAAGAGGCTGAG 1316
Db 1441 CAAAACAGGCTGAATTCAGATTTGAATCATACCAAGAGATGGTTGTAAGAGGCTGAG 1500
QY 1317 GCAGATCTCTGAGAGCAATCAGCAGCTGAACTACTTTAAGAAAGCTCCAAACAGAA 1376
Db 1501 GCAGATCTCTGAGAGCAATCAGCAGCTGAACTACTTTAAGAAAGCTCCAAACAGAA 1560
QY 1377 CAAGCAGCCCAAGGCTCTGAGGAATCTCTGAAATTTATGAGGAGAGAGCTGGTGAAC 1436
Db 1561 CAAGCAGCCCAAGGCTCTGAGGAATCTCTGAAATTTATGAGGAGAGAGCTGGTGAAC 1620
QY 1437 TCGAGAGATTAATCCGATCTGTGAGACAGAGAACTAAGATCAGACATGAACAGAGGA 1496
Db 1621 TCGAGAGATTAATCCGATCTGTGAGACAGAGAACTAAGATCAGACATGAACAGAGGA 1680
QY 1497 AG----- 1498
Db 1681 AGAGATGATTTTCTAGAAAATCACAAACTTGACATTTGTATTAAGCTACTGATTC 1740
QY 1499 -----AGATGATGACACGA 1514
Db 1741 CATTTTGAATATATGTCACCAAAAAACCTGTGTGTTGAAGTGAATGATCACAACGA 1800
QY 1515 CAGGTTTTCATGATTTCAATCAACAGATCCATGAAAGAGAGAGCAAGAGAGAGAA 1574
Db 1801 CAGGTTTTCATGATTTCAATCAACAGATCCATGAAAGAGAGAGCAAGAGAGAGAA 1860
QY 1575 TTTGAGATGTTGACAGAGAGAGAGAGTGCCTGAGGTTGTTGGCAGAGAGAGAGAACT 1634
Db 1861 TTTGAGATGTTGACAGAGAGAGAGAGTGCCTGAGGTTGTTGGCAGAGAGAGAGAACT 1920
QY 1635 TAAATCCCTTAGCAATGACGATTCGGAAG----- 1665
Db 1921 TAAATCCCTTAGCAATGACGATTCGGAAGAGATATATGTAATAATATCCCT 1980
QY 1666 ----- 1665
Db 1981 CTGGCGTTTGTGTTTCAAACTTAAGAGTAATAATTCGGGTTTGTGATCTTTGG 2040

```

```

QY 1666 -AGAGCTGAGAGAGTGTCAAGCTTTCATCGATTTTCAAGAGAAAGAGATGAGAGATTGT 1724
Db 2041 CAGAGCTGAGAGAGTGTCAAGCTTTCATCGATTTTCAAGAGAAAGAGATGAGAGATTGT 2100
QY 1725 GGAAGAGAGGAGATCTCTGATTAAGATCAAGAGAAAGAGATGAGAGATGAGAGAG 1784
Db 2101 GGAAGAGAGGAGATCTCTGATTAAGATCAAGAGAAAGAGATGAGAGATGAGAGAG 2160
QY 1785 GCATCAAGAGAGATTTTGTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAAACAGCTCAT 1844
Db 2161 GCATCAAGAGAGATTTTGTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAAACAGCTCAT 2220
QY 1845 GTACAGCATGGCTTCACATGAGATGATTGA 1878
Db 2221 GTACAGCATGGCTTCACATGAGATGATTGA 2254

RESULT 5
AX078760 3275 bp DNA linear PAT 22-FEB-2001
LOCUS AX078760
DEFINITION Sequence 1 from Patent WO0105951.
ACCESSION AX078760
VERSION AX078760.1 GI:13158379
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Beclin,C., Elmayan,T. and Vaucheret,H.
TITLE Novel gss3 plant gene and use thereof
JOURNAL Patent: WO 0105951-A 1 25-JAN-2001;
AVENTIS CROPS/SCIENCE S.A. (FR) ; INSTITUT NATIONAL DE LA RECHERCHE
AGRONOMIQUE (FR)
FEATURES
source
1..3275
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
primer_bind
693..715
/label="p356AD"
primer_bind
complement(2926)..2952
primer_bind
/label="p356Y"

ORIGIN
Query Match 77.8%; Score 1462; DB 6; Length 3275;
Best local Similarity 83.3%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 0; Indels 376; Gaps 4;
1 ATGAGTTCTTAGGCTGGTCCATGCTTAAGAGAAAGAGCTTCAAGGTGATTAGGCT 60
Db 696 ATGAGTTCTTAGGCTGGTCCATGCTTAAGAGAAAGAGCTTCAAGGTGATTAGGCT 755
QY 61 GAGGTTGAACAGTTGGTTCAAGTTTGGCAGGAGCAGACTGGCTTTCACAAATGAT 120
Db 756 GAGGTTGAACAGTTGGTTCAAGTTTGGCAGGAGCAGACTGGCTTTCACAAATGAT 815
QY 121 GAGAGAGAGTGGAGGTCTATTTCCAAAGAAACAAGAACAAACAGAGAAACTTTGGA 180
Db 816 GAGAGAGAGTGGAGGTCTATTTCCAAAGAGAAACAAGAACAAACAGAGAAACTTTGGA 875
QY 181 AAAAATTTGGTTTCTCAGAATTCGAATCTCTTAAGCTTTGGGGTGTGACAGCAAGGG 240
Db 876 AAAAATTTGGTTTCTCAGAATTCGAATCTCTTAAGCTTTGGGGTGTGACAGCAAGGG 935
QY 241 AAGAGTGAACAGTATCTGGAGAGAGAAACAATGTATCCGAGAGAGTAAACGCAATGCT 300
Db 936 AAGAGTGAACAGTATCTGGAGAGAGAAACAATGTATCCGAGAGAGTAAACGCAATGCT 995
QY 301 CGGGGCAATTCAGATCAATATCTGGTCGGGAGCAGAGCGTTGAGCAAGAAATGATTAAC 360
Db 996 CGGGGCAATTCAGATCAATATCTGGTCGGGAGCAGAGCGTTGAGCAAGAAATGATTAAC 1055

```

QY 361 AACTTGTGGACCCCCCACTGTATCTGCGCTCTTTGGAAAGAGATGGAATTGGCAG 420
 Db 1056 AACTTGTGGACCCCCCACTGTATCTGCGCTCTTTGGAAAGAGATGGAATTGGCAG 1115
 QY 421 GCAGAGAGAGGTTTCTGCTGACACACAGTGTGCGAGGATTTCTGAGGTGAGAGTAT 480
 Db 1116 GCAGAGAGAGGTTTCTGCTGACACACAGTGTGCGAGGATTTCTGAGGTGAGAGTAT 1175
 QY 481 GTGATTAATGCTTCTGAGAGAGATGATTCGATGCTTTGGATGATTTCTGATGACAC 540
 Db 1176 GTGATTAATGCTTCTGAGAGAGATGATTCGATGCTTTGGATGATTTCTGATGACAC 1235
 QY 541 CTTCGAGATGATGATTAATGATCTGCGATGATGATCAAAAAGCCATGATCAAGAAACAG 600
 Db 1236 CTTCGAGATGATGATTAATGATCTGCGATGATGATCAAAAAGCCATGATCAAGAAACAG 1295
 QY 601 AATTAAGTGTGTTCAAAAGTCTTTGGCAGTGTGATGATGCTTGATGATGAGCAGATAAT 660
 Db 1296 AATTAAGTGTGTTCAAAAGTCTTTGGCAGTGTGATGATGCTTGATGATGAGCAGATAAT 1355
 QY 661 GAACCAAGAGAGAGTGTGATGATTCAGAGCTTGTCAAGACGAGCTGTGCTCATGATGAG 720
 Db 1356 GAACCAAGAGAGAGTGTGATGATTCAGAGCTTGTCAAGACGAGCTGTGCTCATGATGAG 1415
 QY 721 TATTAACCTGACACCTCTCTACTAGCTCATGCGAGACAAAAGAGCTAGGAGTTAGCTC 780
 Db 1416 TATTAACCTGACACCTCTCTACTAGCTCATGCGAGACAAAAGAGCTAGGAGTTAGCTC 1475
 QY 781 CATAGAATTTGGCTGAGATTTAGAAAAGATCTACAGATGAGAGCGCATGCTCAT 840
 Db 1476 CATAGAATTTGGCTGAGATTTAGAAAAGATCTACAGATGAGAGCGCATGCTCAT 1535
 QY 841 CCTTGTGTGAGATTTATGCGCAGTGTGAGAGGTTTGGGTGAGATGAAAAGATTTAGAA 900
 Db 1536 CCTTGTGTGAGATTTATGCGCAGTGTGAGAGGTTTGGGTGAGATGAAAAGATTTAGAA 1595
 QY 901 ATTGTCTGGCTCCATATGTCATCATATGATTAATTAAGTGAAGACGATAACGAT 960
 Db 1596 ATTGTCTGGCTCCATATGTCATCATATGATTAATTAAGTGAAGACGATAACGAT 1655
 QY 961 A----- 961
 Db 1656 AAGGTGAATTTCTTCTCTTACTTCTTAATTTTCTCTGCAATTCATGATCTTA 1715
 QY 962 -----AGTGGCTCGCATGTGGCAACCAAGAGCTGTGAAATCTTCCACA 1007
 Db 1716 GAATGTTACATTTGATGCTGCGCATGTGGCAACCAAGAGCTGTGAAATCTTCCACA 1775
 QY 1008 GTATGAGGCTCTTACAGACAGGCACTTCTATGTCACAGAGGCACTGTGGATGAGTGT 1067
 Db 1776 GTATGAGGCTCTTACAGACAGGCACTTCTATGTCACAGAGGCACTGTGGATGAGTGT 1835
 QY 1068 TCTGATTTTGAAGACAGTGTGCACTGCTATTTTGAAGCCGAAACGCTTCAACGGAGTT 1127
 Db 1836 TCTGATTTTGAAGACAGTGTGCACTGCTATTTTGAAGCCGAAACGCTTCAACGGAGTT 1895
 QY 1128 AGCTGAAGTGGGTTAATTAATAATTTGCTGGGTCACAAAGCGCATGATGTTTCTGAGG 1187
 Db 1896 AGCTGAAGTGGGTTAATTAATAATTTGCTGGGTCACAAAGCGCATGATGTTTCTGAGG 1955
 QY 1188 TGTTCGCAACTGTATGCTCTCTTGCACAGACAGATGTGACATATTCATCAACA 1247
 Db 1956 TGTTCGCAACTGTATGCTCTCTTGCACAGACAGATGTGACATATTCATCAACA 2015
 QY 1248 CTCTCA----- 1253
 Db 2016 CTCTCAAGGTTCTCTCCCCCAAGAAATTTGATATATCTTTAGTTTGTCTATGGAAT 2075
 QY 1254 -----AGG 1256
 Db 2076 TTAAGTTTGTGTGCTGCTGTATATGATCTGTATATATATCTATGATTCATTTAGG 2135

QY 1257 CAAAACAAGCTGAATTCGAGTTGAATATACCAAGAGTGTGTTAAAGAGCTGAG 1316
 Db 2136 CAAAACAAGCTGAATTCGAGTTGAATATACCAAGAGTGTGTTAAAGAGCTGAG 2195
 QY 1317 GCAGATCTTGAAGACAAATCAGAGCTGAACTACTTTAAGAACAGCTCTCAAAACGAA 1376
 Db 2196 GCAGATCTTGAAGACAAATCAGAGCTGAACTACTTTAAGAACAGCTCTCAAAACGAA 2255
 QY 1377 CAACGACCCCAAGGCTTGAAGATCTCTGGAATTTATGAGCAGAGAGCTGTAGAAC 1436
 Db 2256 CAACGACCCCAAGGCTTGAAGATCTCTGGAATTTATGAGCAGAGAGCTGTAGAAC 2315
 QY 1437 TGCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
 Db 2316 TGCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2375
 QY 1497 AG----- 1498
 Db 2376 AGAGGTATGATTTTCTTGAAGAAATCAACAACTTGACATTTTGTATTACTTACTGATTTCA 2435
 QY 1499 -----AGATGATGACACGA 1514
 Db 2436 CATTTTGAATTAATTTCTCAACAAACAAACCTGTGTGTTTGAAGATGATGACACAG 2495
 QY 1515 CAGTTTTCATGATGATTCATCAACAGATTCATGAAAGAGAGAGAGAGAGAGAGAGAG 1574
 Db 2496 CAGTTTTCATGATGATTCATCAACAGATTCATGAAAGAGAGAGAGAGAGAGAGAGAG 2555
 QY 1575 TTTGAGATGTTGCGAGCAG 1634
 Db 2556 TTTGAGATGTTGCGAGCAG 2615
 QY 1635 TAATCCCTCTAGCAATGACGATGCGGAAAG----- 1665
 Db 2616 TAATCCCTCTAGCAATGACGATGCGGAAAGAGATTAATGATTAATTAATTAATTCCT 2675
 QY 1666 ----- 1665
 Db 2676 CTGGCGTTTGTGTTTCAAACTTAAGATTAATTAATTAATTAATTAATTAATTAATTCG 2735
 QY 1666 -AGAGTGAAGAGAGTCAAGCTTCAAGATTTCAAGAGAGAGAGAGAGAGAGAGAGAG 1724
 Db 2736 CAGAGCTGAAGAGAGTCAAGCTTCAAGATTTCAAGAGAGAGAGAGAGAGAGAGAG 2795
 QY 1725 GGAAGAGAGAGAGAGTCAAGATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1784
 Db 2796 GGAAGAGAGAGAGAGTCAAGATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2855
 QY 1785 GCATCAAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1844
 Db 2856 GCATCAAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2915
 QY 1845 GTACAAAGCATGGCTTCAACATGAAGATGATGA 1878
 Db 2916 GTACAAAGCATGGCTTCAACATGAAGATGATGA 2949

RESULT 6
 AB025633 81365 bp DNA linear PIN 27-DEC-2000
 LOCUS Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MGM1.
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MGM1.
 ACCESSION AB025633 Ba000015
 VERSION AB025633.2 GI:10178221
 KEYWORDS
 SOURCE
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases)
 REFERENCE Sato,S., Nakamura,Y., Kaneko,T., Katoh,T., Asamizu,E., Kotani,H.
 and Tabata,S. Structural analysis of Arabidopsis thaliana chromosome 5. X.

Sequence features of the regions of 3,076,755 bp covered by sixty P1 and TAC clones
DNA Res. 7 (1), 31-63 (2000)

JOURNAL
MEDLINE
PUBMED
20181125
10718197
2 (bases 1 to 81365)
Nakamura, Y.
Direct Submission
Submitted (02-APR-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamura@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)
On Sep 15, 2000 this sequence version replaced gi:4589439.
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see <http://www.kazusa.or.jp/kaos/cgi-bin/sgd/graph.cgi?c=MQM1>
Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/Grail-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and Slicepredictor (Volker Brendel, Stanford University, <http://gremlini.zool.iastate.edu/cgi-bin/bp.cgi>).
Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).
This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K19M13 and the 3' clone is M0011.

FEATURES
source
location/Qualifiers
1..81365
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MQM1"
/clone_1fb="Mitsui P1"
join(3184..3403,3474..3595,3681..3723,3777..3898,4042..4198,4445..4515,4601..4714,4888..4988,5300..5420)
/note="gbl|AF01580.1"
gene_id:MQM1.24
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BA97251.1"
/db_xref="GI:88097210"
/translation="MHVQPGTSDVDEPAHKKVDTPPKHVOLEPIPSKTVVDNAQNS YSAGFDPSSNSPLSPVSESSSEEGDDPLPGIEDLOISEPYGHLQAG YSNGTSCNFEWCHLSDGSVNYIDESSRERTSNAGKQPNYLVADVDLYLAEV QLPDDNRKRGKRAFGKYRDLTVLADHVLFSADCFSVFPMFAIFCYLLIGG ACQDDPMOSNIETKLTGHASVYSLAVGPDVWBAATLSIKREGYSIKCISDITLAE KESASTVTITPGQPAELVITGSDSEHSIRANSGPDLIGSRDEIVLTIRLEIKVM FFAANSFYHFRSCSKSRGNTL"
join(6317..6428,6508..6629,6714..6756,6841..6932,7121..7169,7396..7466,7550..7663,7785..7882,8112..8181)
/note="gbl|AF01580.1"
gene_id:MQM1.23
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BA97250.1"
/db_xref="GI:8809709"
/translation="MQGTSVSEPALNQDETPPISTKTVVNDTKNIFFYDDGDAPL PALNETOISEPYGHLQAGYNSNGTSCNFEWCHLSDGSVNYIDAKKKRNYVT ADVUGCLAIEVQPLDDNRKRGKLVAVFANDNRKNIQHPEDMSNIIDKTLTGHASVY SLAIGFVHTEAATLSIREGTYTKCNDLITTEKKSASTAVAKIPEKPELVITGSD GSHCLRVNDSEWDIRLCKEERVLVKTVVVEQL"
complement(join(8642..8705,8794..8863,8978..9387,

9584..10347))
/note="gene_id:MQM1.22
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BA97249.1"
/db_xref="GI:8809708"
/translation="MSMKKSSGWTAPDLKOROKOGLSEVEGDPFPVSTSVNASF GVRGRLRHHEPEKSESSVLLPPSRPALTEKDCGNGERGCGCKRKPDTLSIPVNS HDLAFKLKEMNSWADNLTIRVLTSTEDDFEMALAFKGMSSGGEDEPTSKIEGY SSNRSSEYTRFEKTYTSSVKAAKSTFEDAGKTYLJENSOGSFLVNASDNEKFPDI SELDSTLIQLOSTPIPEMEEDLYSHRKDALKWRASASNHRAQNAFORDHASA KOHSDVARDEWLAELKNAEAKKIIGITKNDKNIWKLDLHGHLATEAVQALCKEOM IEGHFVNSVSVNSRGRSKAALARSASQEPFGLEDGEGCHQCGSSRELRNSLOVITG IGHSHRQSLPIAVKTFPFDNRYPREDRPPGVTITRPPFRHS"
complement(12417..13424)
/note="contains similarity to unknown protein
gbl|AF27018.1
gene_id:MQM1.21"
/codon_start=1
/evidence=not experimental
/protein_id="BA97248.1"
/db_xref="GI:8809707"
/translation="NATDSQPMQKLTLPKTRIALTVISTMTDMAQRPDGTINRFLR LEPFAPAPPKPVNIIVTSDPVDOSDILMFRLYTPHVGDKIPVVPFGGAFALS PNAVPYDNCRRPARKLPAYVISVNYRLAEPHRYPOYDGFALXYIEHNSILPA NADLSRCFAGSAGCNIAHVAIRICREPRSFTVKLIGLISIOFPFGERTAB KOLVGAPLVPDRTDWCAMGLNDRHEVNVNGPNAVVISGLDIPEYTVVYAGPPL KMQRSYVWMLCKGKATLIEPNNFHFVLPPELPEAGQLIMRIKDVDEKVASLS A"
join(14922..14981,15072..15167,15563..15666,15755..15891,16096..16178)
/note="gene_id:MQM1.20"
/codon_start=1
/evidence=not experimental
/product="50S ribosomal protein L24"
/protein_id="BA97247.1"
/db_xref="GI:8809706"
/translation="MGWKAELIRHMKTLRGDNWMIIRGKDGKGTIKRIVRSQNR VYEGKNLIKHIKGGPDHIEGCIFFTEAPLHASNVOYDVPYGRPKGVKYLIEDTK VVARGTGSGSIIPEPBLIKRATPRPTAGKQDPMFSEWQITDAGTKMPUL"
join(16823..16952,17143..17320,17412..17493,17675..17851,17943..18110,18199..18390)
/note="gene_id:MQM1.19"
/codon_start=1
/evidence=not experimental
/product="26S proteasome, non-ATPase regulatory subunit"
/protein_id="BA97246.1"
/db_xref="GI:8809705"
/translation="MERLQRIFGAGGIGHASPDSPTLTDSBOVYISLALIKMLKHG RAGVPMHNGMLGERTVYTRVYDVPRAMPDSGTVSEANDHYQTMKMLDKOTG RBEMVGVYHSHRGECWLSGVINTQSFALNQAVAVVVDPIOSVGVKVIIDFR SINPQITMIGQBPQRTSNLGHINRKSIALHGLNRHYSLAINRNKLELVEKNN LHKKWTIDTLTRRPFTHSKTNGQVOEMLSLAAKYNKAVOEDSELPEKLAIVMNGR OAKKGLSEHVNLSNNSNVTOTIGTMLDTPV"
complement(join(18965..19043,19240..19298,19419..19536,19666..19739,19908..19961,20387..20530))
/note="contains similarity to unknown protein
emb|CAA22897.1
gene_id:MQM1.18"
/codon_start=1
/evidence=not experimental
/protein_id="BA97245.1"
/db_xref="GI:8809704"
/translation="MDKMQALRKMKLVGMEVEEQQADESSLSFEMEDLNRCAL TKQRFYPAICLSAGLTCTLSMLVFNPVFGITFTGNLMALGSTFALIGPOKOV TMMLPARIYATYALVYLAIIIFCALVYRNKLTLLAIIIEFTGLIWSLSIIPAR TWXSLFMCVPTFEF"
join(12262..12363,123708..123999,24111..24355,24458..24624,24715..24925)
/note="emb|CAB62356.1
gene_id:MQM1.17
similar to unknown protein"

CDS

CDS

CDS

FEATURES
source

CDS

CDS

CDS

CDS

CDS


```

/codon_start=1
/evidence=not experimental
/protein_id="BA97244.1"
/db_xref="GI:8809703"
/translation="MSSRAGPMSKERKYNQGYRPEVEVDVGLACTRLASDDGGEG
EYISKKNKPKPNTSGKTWVSQNSNPPRMWGQOGRSNVSGRNNVSGKNGRG
IOANTSGRGRALSRKYDNFVAPPPVSPLEGGMWOARQSAQTLVQEPEDD
YNSASPEEDSDALDDSDDDLDSDYDVSQSHSGSKQNMFKFPGSLDSIEQ
INBPORHICPAQONCPRAIDWYNLHPLHARTKARGLHRELAELVEDLOMRG
ASVTPGCBTGYGKGLGDEKDEKDEYELWPMVIMNTRLDKDNMDKMLGMNGOLEYF
DKKRALRARSYGPQGRMSVLMPESSATGYLEBRHLRELAELGDIRIAMGKRSV
FSGGVOLYGLATKODLDIENQSGKTRLFELKSIOEMVXKELRQISENQDLY
FKKLSKONKHAIVLESLIEMSEKLRTAEDNRIYRQSTKQNMKNEBMDHFRF
MSIKOIHERRDAKEENFEMLOOQERAKVGOQOONINSNDRCRAEVSFIEF
QKEMEFYBERHEMLIKDEKDEKMKHHEIIPDLKEPDEALEQJLMYKGLHNE
D"
join(25614..25838,26061..26207,26324..26438,26516..26610,
26710..26813,27522..27765,27943..28036,28115..28305,
28385..28604,28688..28855,28944..29092)
/notes="gb|AAC97420.1
gene_id:MOM1.16
strong similarity to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BA97243.1"
/db_xref="GI:8809702"
/translation="MAPPAQOTAAVAEAAGADGAPPOQRGFGSTISGIVRAVWFY
ASKFSPKQKMDPTASQMTNLPHKESLDWMFYLSEQEKENDGNRRLYMBETN
IPAVWTPBSIRTKSLIYVPSSELPONNGSLYAHIPARSGFPIPDPEOPILMSFR
THAVATFPOKKNKKSLLSGSKPDSPEVEKDGKSDKEEVPWEMISLMKN
VTINLVDFQPADLVEPTTGNTPTTYNFMWMLRDKRIPNETHSELPLNLEISF
ISMWKQDLPQVQDSFQMRQSYSGMLGSDDELKFLPBLGITHFMSLSHVF
DPLAFKNDIOFMNKNKSMBSAKSVLANICQFVLFYLDNDTSMILASGVVC
IEFWKIKAMRIEVDRSAMIPLRFLHREBSYANIKEXYDIAIKELSVLLLVGL
SYSLAYERRHSWYSWISLSITCVYFIMMCPOLFVYKLSVAHLPWMTYK
LNTIIDLPFAVFMPLIHLRHSVPRDVIPLIYQWVYPVDTKRVNREGFGEDET
AEKLLTEKEBEDKTN"
complement(join(29439..29576,29661..29885,29985..30268,
30341..30493,30579..30722,30956..31464))
/notes="gene_id:MOM1.15
77.8%; Score 1462; DB 8; Length 81365;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 0; Indels 376; Gaps 4;

```

```

QY 421 GCAGAGAGAGGTTCTGCTGACGACACAGCGTGTGACAGATTCTGACGTGAGATGAT 480
DB 23092 GCAAGAGAGAGGTTCTGCTGACGACACAGCGTGTGACAGATTCTGACGTGAGATGAT 23151
QY 481 GTGGATTAATGCTTCTGAGAGAGAAATGATTCGATGCTTTGGATGATTTCTGATGCGAC 540
DB 23152 GTGATTAATGCTTCTGAGAGAGAAATGATTCGATGCTTTGGATGATTTCTGATGCGAC 23211
QY 541 CTTCGAAGTATGATTAATGATCTGAGATGAGTCAAAAGAGCCATGATCAAGAAAGAG 600
DB 23212 CTTCGAAGTATGATTAATGATCTGAGATGAGTCAAAAGAGCCATGATCAAGAAAGAG 23271
QY 601 AATAAGTGTTCAAAAAGTCTTTGGACGCTGATAGCTTGTGATCGACGATTAAT 660
DB 23272 AATAAGTGTTCAAAAAGTCTTTGGACGCTTGGATAGCTTGTGATCGACGATTAAT 23331
QY 661 GAACCAAGAGGCGATGCTGATGCTTGTGACGACGAAAGAGCTAGCGAGTTAAGCTC 720
DB 23332 GAACCAAGAGGCGATGCTGATGCTTGTGACGACGAAAGAGCTAGCGAGTTAAGCTC 23391
QY 721 TATAACCTGCAACCTCTAATAGCTCATGCGAGGCAAAAGAGCTAGCGAGTTAAGCTC 780
DB 23392 TATAACCTGCAACCTCTAATAGCTCATGCGAGGCAAAAGAGCTAGCGAGTTAAGCTC 23451
QY 781 CATAGAGATTTGCTGAAGTTTGAAGAGATCTACAGATGAGAGCGCATCTGTCAAT 840
DB 23452 CATAGAGATTTGCTGAAGTTTGAAGAGATCTACAGATGAGAGCGCATCTGTCAAT 23511
QY 841 CCTGTGTGAGATTTATAGGCACTGAGAGGTTTGGGTGAGATGAAGAGATTAAGAA 900
DB 23512 CCTGTGTGAGATTTATAGGCACTGAGAGGTTTGGGTGAGATGAAGAGATTAAGAA 23571
QY 901 ATTGTCTGAGCTCAATGATCATCATGATTAATCTAGACTGATTAAGAGATTAAGAT 960
DB 23572 ATTGTCTGAGCTCAATGATCATCATGATTAATCTAGACTGATTAAGAGATTAAGAT 23631
QY 961 A----- 961
DB 23632 AAGGTGAATCTCTGCTCTTTTACTTTTAAATTTTCTCTTGCAATCTGATCTTA 23691
QY 962 -----AGTGGCTGGCATGGGCAACCAAGAGCTGTGAATATCTTGACAA 1007
DB 23692 GAATGTACATTTGATGCTGCTGCAATGGGCAACCAAGAGCTGTGAATATCTTGACAA 23751
QY 1008 GTATGAGGCTCTTAAGACGCGCATTTCTATGCTCAAGGCGCATCTGGGATGAGT 1067
DB 23752 GTATGAGGCTCTTAAGACGCGCATTTCTATGCTCAAGGCGCATCTGGGATGAGT 23811
QY 1068 TCTGATGTTGAGAGAGTGCACATGCGATTTGAGAGGCGCAACGCTCCACCGGAGTT 1127
DB 23812 TCTGATGTTGAGAGAGTGCACATGCGATTTGAGAGGCGCAACGCTCCACCGGAGTT 23871
QY 1128 AGCTGAGTGGGTTGATAGATTTGCTGGGCTCAAGAACGCAATGTTTCTGAGG 1187
DB 23872 AGCTGAGTGGGTTGATAGATTTGCTGGGCTCAAGAACGCAATGTTTCTGAGG 23931
QY 1188 TGTTCGCAACTGATGCTTCTTCAACGAAGCAAGATCTGGAATTTCAATCAACA 1247
DB 23932 TGTTCGCAACTGATGCTTCTTCAACGAAGCAAGATCTGGAATTTCAATCAACA 23991
QY 1248 CTCTCA----- 1253
DB 23992 CTCTCAAGTCTCTCCCAAGAAATTTGATATATGCTTTAGTTTGTCAATGAAT 24051
QY 1254 -----AGG 1256
DB 24052 TTAAGTTTGTTCGCGTTAATGATCTGTTATGATATATATGATTAATGATTAAG 24111
QY 1257 CAATAAGAGCTGAATTTGAGATTTGAATCATCAAGAGATGTTGTAAGAGCTGAG 1316
DB 24112 CAATAAGAGCTGAATTTGAGATTTGAATCATCAAGAGATGTTGTAAGAGCTGAG 24171
QY 1317 GCAGATCTGAGAGCAATGACGAGCTGAATCTTTAAGAACAGCTCTCAAAAACAGAA 1376

```



```

Db 24172 GCAGATCTGAGAGACATCAGACGCTGAACTTTAAGAAACAGCTCAAAACAGAA 24231
QY 1377 CAGACGCCAAGGTGCTTGAAGAAATCTGGAATATTATAGAGAGAGAGTGGCTGAAC 1436
Db 24232 CAGACGCCAAGGTGCTTGAAGAAATCTGGAATATTATAGAGAGAGAGTGGCTGAAC 24291
QY 1437 TGCAGAGATTAATCGATCGTGAAGACAGAGAACTTAAGATCAGCATGAACAGAGGGA 1496
Db 24292 TGCAGAGATTAATCGATCGTGAAGACAGAGAACTTAAGATCAGCATGAACAGAGGGA 24351
QY 1497 AG----- 1498
Db 24352 AGAGATGATATTTTCTAGAAATCACAAATTGACATTTGTATTACTTACTGATTCA 24411
QY 1499 -----AGATGATCAGACCGA 1514
Db 24412 CATTTTGAATTAATTTGTCCAAACAAAACCTGTGTGTGTTGAAGATGATCAGACGA 24471
QY 1515 CAGGTTTTCATGATTCATCAATCAAAACAGATTCATGAAAGAGAGACGCAAGAGAGAGAA 1574
Db 24472 CAGGTTTTCATGATTCATCAATCAAAACAGATTCATGAAAGAGAGACGCAAGAGAGAA 24531
QY 1575 TTTTGAATGTTTSCAGACGAGAGAGACGTCGCAAGTTTGGCCAGACGAGAGAGAAAT 1634
Db 24532 TTTTGAATGTTTSCAGACGAGAGAGACGTCGCAAGTTTGGCCAGACGAGAGAAAT 24591
QY 1635 TAATCCCTCTAGCAATGACATGTCGCAAG----- 1665
Db 24592 TAATCCCTCTAGCAATGACATGTCGCAAGAGATTAATGATTAATCAATTAATCCCT 24651
QY 1666 ----- 1665
Db 24652 CTGGCGTTTGTGTTTTCCTTCAAACTAAGATGATTAATTCGCGTTTGTGTTTCG 24711
QY 1666 -AGAGCTGAGAAAGTGTCAAGCTTCATGAGTTTCAAGAGAAAGATGAGAGAGTTGT 1724
Db 24712 CAGAGCTGAGAAAGTGTCAAGCTTCATGAGTTTCAAGAGAAAGATGAGAGAGTTGT 24771
QY 1725 GGAAGAGAGAGAGATGCTGATTAAGATCAAGAGAGAGAGATGAGAGACATGAAGAGAG 1784
Db 24772 GGAAGAGAGAGAGATGCTGATTAAGATCAAGAGAGAGAGATGAGAGACATGAAGAGAG 24831
QY 1785 GCATCAGAGAGATTAATTTGATCTGAGAGAAAGATTTGATGAGGCTTTGAGACAGCTCAT 1844
Db 24832 GCATCAGAGAGATTAATTTGATCTGAGAGAAAGATTTGATGAGGCTTTGAGACAGCTCAT 24891
QY 1845 GTACAGCATGAGCTTCACATGAGAGATGATTGA 1878
Db 24892 GTACAGCATGAGCTTCACATGAGAGATGATTGA 24925

```

```

RESULT 7
ATH528171 650 bp DNA linear PLN 29-MAR-2003
LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone
DEFINITION 157C08
ACCESSION AJ528171
VERSION AJ528171.1 GI:26796431
KEYWORDS left border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eustosils II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F.,
Chauvin, S., Bechtold, N., Crandall, C., Depose, R., Pelletier, G.,
Leplintec, L., Caboche, M. and Lecharny, A.
T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
JOURNAL EMO Rep. 3 (12), 1152-1157 (2002)
MEDLINE 22363535

```

```

PUBMED 1244565
REFERENCE 2 (bases 1 to 650)
AUTHORS Balzerque, S.
TITLE Direct Submission
JOURNAL Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (http://www.genoplante.com and
http://genoplante-info.inbio.inra.fr).
FEATURES
Source
1..650
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Massillaeskija"
/db_xref="taxon:3702"
/clone="157C08"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
1..650
/notes="T-DNA flanking sequence
left border"
ORIGIN
Query Match 23.3%; Score 437.8; DB 8; Length 650;
Best Local Similarity 87.4%; Pred. No. 9e-96;
Matches 522; Conservative 0; Mismatches 2; Indels 73; Gaps 1;
QY 733 CCTTACTAGCTCATGCGAGACAAAGAGAGCTAGGCGAGTTAAGTCCATAGAGAAATTG 792
Db 1 CCTTACTAGCTCATGCGAGACAAAGAGAGCTAGGCGAGTTAAGTCCATAGAGAAATTG 60
QY 793 GCTGAAGTTTGAAGAAAGATCTACAGATGAGAGGCGCATCTGCTTCTGGTGGAG 852
Db 61 GCTGAAGTTTGAAGAAAGATCTACAGATGAGAGGCGCATCTGCTTCTGGTGGAG 120
QY 853 ATTTATGCGCAGTGAAGGTTTGGTGAAGATGAAGAAAGATTAATGATGCTGCGCT 912
Db 121 ATTTATGCGCAGTGAAGGTTTGGTGAAGATGAAGAAAGATTAATGATGCTGCGCT 180
QY 913 CCAATGTCATCATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 961
Db 181 CCAATGTCATCATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
QY 962 ----- 961
Db 241 TTCTGCTTTTACTTCTTAATTTTCTTGTGCAATCTAATCTTAATTAATTAATTAATTAAT 300
QY 962 --AGTGGCTCGGATGAGCAACCAAGAGCTGTGATACTTTCAGACAGATGAGGCTCT 1019
Db 301 GTATGCTCGGATGAGCAACCAAGAGCTGTGATACTTTCAGACAGATGAGGCTCT 360
QY 1020 TAGAGCAGCCATCTCTATGTTCCACAGGCGCATGCTGAGATGAGTGTCTGATGTTGA 1079
Db 361 TAGAGCAGCCATCTCTATGTTCCACAGGCGCATGCTGAGATGAGTGTCTGATGTTGA 420
QY 1080 GAGCAATGCACTGGGTATTTGAGAGCGCAAGCGCTCCACCGGAGATTAAGTGAATGGG 1139
Db 421 GAGCAATGCACTGGGTATTTGAGAGCGCAAGCGCTCCACCGGAGATTAAGTGAATGGG 480
QY 1140 GTTAGATTAATTTGCTGGGGTCAAGAGCGCATGATGTTTCTGAGGTGTTCCCACT 1199
Db 481 GTTAGATTAATTTGCTGGGGTCAAGAGCGCATGATGTTTCTGAGGTGTTCCCACT 540
QY 1200 GTATGCTTCTTTCAGACAGAGAGATGTCATATTAATCAACTCTCAAG 1256
Db 541 GTATGCTTCTTTCAGACAGAGAGATGTCATATTAATCAACTCTCAAG 597

```


QY 1057 GGATGAGTGTTCGATGTTTGAGAGACAGTGCCTGATTTTGAGAGCCGACCTC 1116
 |||||
 Db 1070 GCGATGAGTGTCTGATCTTTTGACAGCTCGCTGTGGCTATATGAGAGAGAACTTTG 1129
 |||||
 QY 1117 CACCGGAGTGTGAGATGAGGAGTTAGATGATGCTCTGGGT-----CAGAACCG 1169
 |||||
 Db 1130 CATGATCATCTTGTTCGTCAAAAGACAGACAGAAATCATGAAACAGTGTCTACCAAGGT 1189
 |||||
 QY 1170 CATATGTTTCTGAGAGTGTTCGCCAATGATGCTTCTGCAACGAGCAAGATCT 1229
 |||||
 Db 1190 TACATTTTACCTGCTGGGAGAAAGCAGCTGTATGTTTCTTGCTCAAAAGATGACAT 1249
 |||||
 QY 1230 GGCATATTCATCAACACTCTCAGGCAAAACAGGCTGAATTCAGTTGAATCAT 1289
 |||||
 Db 1250 GAGACATTCATATGACATTTGCCACGAGAAATCCCGTTGAATATGATGATCTTTA 1309
 |||||
 QY 1290 CCAAGAGATGTTGTAAAGAGCTGAGGACGATCTCTGAGACATCAGACGCTGAACT 1349
 |||||
 Db 1310 TAATGAGATGTGTGTTTACCAATGAAAGCAGATGAGGAGAACACAGCACTGAATTA 1369
 |||||
 QY 1350 CTTTAAAGACAGCTCTCAAAACAGAACAGCAGCAGAGTGTGAGAACTCTCGA 1409
 |||||
 Db 1370 CCTGAAGAACAAATGTTTAAAGAGCAGCACTTAAATTTAGTAGAGATCATTTGAG 1429
 |||||
 QY 1410 AATTATGAGGAGAGAGTGGCTGAACCTGCAAGGATATCGGATCGTAGACAGAGAAC 1469
 |||||
 Db 1430 TGTGTTTACCAAAACTTTCGAGACATGAGAGAAATATCTTTTACAAATTAAGC 1489
 |||||
 QY 1470 TAAATGACAGCATGAAACAGAACAGGAGATGATGACACAGACAGTCTTTTCATGA 1529
 |||||
 Db 1490 CAAAGAAAGCACTTGAGATGATGAAAGAGATGAAAGTACAGAGAGAAATTTTTCATGA 1549
 |||||
 QY 1530 TTCAATC-AAACGATTCATGAAAGAGACGCAAGAGAGAAATTTGAGATGTTGC 1588
 |||||
 Db 1550 TCAATTCGAAAAAATTCACAGAGCCACAGAGAGAAAGAAATCAAGTTTGAAGAAAGTTGC 1609
 |||||
 QY 1589 AGCAGACAGAGACGTCGCAAGGTTGTTGGCAGAGCAGACATTAATCCCTTAGCA 1648
 |||||
 Db 1610 TCGAAGAGGAGCAGACAAAGGCT-----AGACAAAGTATGTTGACTCTGGAAGTA 1660
 |||||
 QY 1649 ATGACGATTCGCGAAAGAGAGCTGAGAGAGTGTCAAGCTTCATGATTCAGAGAGAAAG 1708
 |||||
 Db 1661 CTGAAGATGCGAGGCAAAAGAGAGAAAGATTCAGAACTTCATGACTGCGAGGTGAAG 1720
 |||||
 QY 1709 AGATGAGAGATTTTGGAAGAGAGGAGATGCTGATAAAGATCAAGAGAGAAAGATGG 1768
 |||||
 Db 1721 ATGTTGAGGAATTTGAAGCCGAAACGAGACAACTGATCAATTTGCATGAAAGAGAAAGG 1780
 |||||
 QY 1769 AAGACATGAAAGAGAGGATCAGAGAGATTTTGTCTGGAAGAAAGATTTGATGAGG 1828
 |||||
 Db 1781 TGAAGCTCAAGAGAGAGACTCTGCGCAGAGAAATTTGAGCTTGAGAGAGAGACTGACACAG 1840
 |||||
 QY 1829 CTTTGAAGACGCTCATGTACAGCA 1853
 |||||
 Db 1841 CCCTCACTTCCCTGATGATTAACA 1865
 |||||

RESULT 9
 AKI00699 2570 bp mRNA linear PLN 24-JUL-2003
 LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J023114M02, full
 DEFINITION
 insert sequence.
 ACCESSION AKI00699
 VERSION AKI00699.1 GI:32985908
 KEYWORDS
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriatoidae; Oryzaceae; Oryza.
 REFERENCE 1
 AUTHORS The Rice Full-length cDNA Consortium, National Institute of

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 2 (bases 1 to 2570)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imanura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirizane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
 Namiki, T., Narikawa, R., Nikura, J., Nishi, K., Nomura, K.,
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
 Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and
 Yoshimura, A.
 Direct Submision
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
 Agrobiological Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail:skikuchi@ias.affrc.go.jp,
 Tel:81-29-838-7007, Fax:81-29-838-7007)
 This clone is one of the 28k full-length cDNA clones from japonica
 rice. URL: <http://cdna01.dna.affrc.go.jp/cDNA/>
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and
 Yamamoto, M.
 FALS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
 Mizuno, K., Narikawa, R., Nikura, J., Oka, M., Ryu, R., Sugano, S.,
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
 Yoshimura, A., Matsubara, K., and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, F., Iida, J., Imanura, K., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
 Kishikawa-Hirizane, T., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,

TITLE JOURNAL

COMMENT

Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1. 2570
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023114M02"

ORIGIN

Query Match 19.8%; Score 372; DB 8; Length 2570;
Best Local Similarity 57.0%; Pred. No. 1e-79;
Matches 786; Conservative 0; Mismatches 575; Indels 19; Gaps 5;

477 TGATGTGATTAATCTCTGAGGAGAGAAATTCGATGCTTGTGATGATTTGATGA 536
Db 713 TGAATGACAAATGATGAGATTAATGATCATATGACGATTCGATGATGATGAT 772
Qy 537 CGACCTTGCAAGATGATTAATGATCGATGATGATCAAGAGCATGATCAAGAA 596
Db 773 TGATGATACCAAGTGTGATCTTGAATCTTAAGGCAATCGAAGAAATTTGAGCTGAAA 832
Qy 597 GCAGATTAAGGTGTTCAAAAGTTCTTGGCAGCTGATAGTTCATGATGATGAT 656
Db 833 AAGGACAAAGCTGTGAAAGCAATTTTGAACCTTGGAGAGTTAAGTGTGAACAGAT 892
Qy 657 AAATGAAACAGAGAGAGAGATGTCATGTCAGCTTGTTCAGAACGAGACTGTGCAATCA 716
Db 893 AAATGAAAGAGCTAGGCAATGTCATGTCAGCAATG-CAAAATGATGATGATGAT 951
Qy 717 TTGTATTA---CCTGCAACCTCTACTACTGATGATGATGATGATGATGAT 773
Db 952 CTGTGACAAAGGTTAAGCTTGTGATGATGATGATGATGATGATGATGAT 1011
Qy 774 TAAAGCTCAATAGAGATTTGGCTGAAGTTTGAAGAAAGATTTAAGATGAGAGCGATC 833
Db 1012 TAAAGCTCAATAGAGATTTGGCTGAAGTTTGAAGAAAGATTTAAGATGAGAGCGATC 1071
Qy 834 TGTCTATCTTGTGTGATGATTTAAGGCAATGAGAGGTTTGGTGTGATGATGAAAGCA 893
Db 1072 AATGATCATCTGCGAGATTTGTTGATGATGATGATGATGATGATGATGAT 1128
Qy 894 TTAATGAAATTTGCTGGCTCTCAATGATGATGATGATGATGATGATGATGAT 953
Db 1129 TGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
Qy 954 TAAAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1013
Db 1189 GATGATCAAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1248
Qy 1014 GGTCTTAAGGAGAGAGATTTCTATGATGATGATGATGATGATGATGATGAT 1073
Db 1249 CGTGAAGAGAGAGATTTCTATGATGATGATGATGATGATGATGATGATGAT 1308
Qy 1074 GTTGAAGAGAGAGATTTCTATGATGATGATGATGATGATGATGATGATGAT 1133
Db 1309 ATTGAAGAGATTTCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1368
Qy 1134 GATGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
Db 1369 TCAAGAGAGAGAGATTTCTATGATGATGATGATGATGATGATGATGATGAT 1428
Qy 1194 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
Db 1429 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1485
Qy 1254 AGGCAAGAGAGAGATTTCTATGATGATGATGATGATGATGATGATGATGAT 1313
Db 1486 TGAAGAGAGATTTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 1545

Qy 1314 GAGCGATCTCTGAGGCAATCAGAGCTGATGATGATGATGATGATGATGATGAT 1373
Db 1546 GAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1605
Qy 1374 GAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1433
Db 1606 GAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665
Qy 1434 AACTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493
Db 1666 AGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1725
Qy 1494 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553
Db 1726 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1785
Qy 1554 AAGAGAGAGAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1613
Db 1786 CACGAGAGAGAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1845
Qy 1614 TGCCAGAGAGAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1673
Db 1846 AG-----AGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1896
Qy 1674 GGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1733
Db 1897 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1956
Qy 1734 GGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1793
Db 1957 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2016
Qy 1794 GGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1853
Db 2017 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2076

RESULT 10
AK064217
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:002-104-F07, full
insert sequence.

ACCESSION
AK064217
VERSION
AK064217.1 GI:32974235
KEYWORDS
FLI CDNA; oligo-capping.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriaristidae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team,
Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Oneda, E., Yahagi, M., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Oono, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurose, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishida, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Minura, J.,
Kunemitsu, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Hayatsu, N., Imoto, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, M., Hayatsu, N., Imoto, K., Arakawa, T., Fukuda, S.,
Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Otsu, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
JOURNAL
Science 301 (5631), 376-379 (2003)
MEDLINE
22752273
PUBMED
12869764

RESULT 11
AF542974 1564 bp mRNA linear PLN 02-JUL-2003
LOCUS Triticum aestivum Emr1 mRNA, complete cds.
DEFINITION AF542974
ACCESSION AF542974.1 GI:32401385
VERSION
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Triticum.
1 (bases 1 to 1564)
Zhao, X., Li, Q. and Zhang, X.
Isolation and expression of a new kind of gene involve in
embryogenesis in Triticum aestivum L.
Unpublished
2 (bases 1 to 1564)
Zhao, X., Li, Q. and Zhang, X.
Direct Submission
Submitted (03-SEP-2002) College of Life Sciences, Shandong
Agricultural University, Daizong Street, Tai'an, Shandong 271018,
P. R. China
Location/Qualifiers
1. 1564
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultiivar="PH 82-2-2"
/db_xref="taxon:4565"
/dev_stage="embryo"
309..1223
/note="involved in embryogenesis"
/codon_start=1
/product="Emr1"
/protein_id="AAP80862.1"
/db_xref="GI:32401386"
/translation="MGQELLEYSDYATKARAYPGGHRGMSVLIPESSAVGWE
AERLKHPIIDRTDIDONRVRPFLPGKQLTGFARKEDETHFRHOGKSPKX
EMRSHNEVVAOMKQSEEDNQNLVLRKRVKRSKVEETLGVITQKLRMBEN
IIVRSKAKKSEVEEEMKSEKPFHDIENIHATBEDKSEFRLQDEERARQCD
VDSGTEENRRIRKQVORFICQVADVQEFPAERDEMIKAEKVKVQLKEVAKVEE
LEKFDALVTGLMEKRRPRTFOASSSP"

ORIGIN
Query Match 15.5%; Score 291.6; DB 8; Length 1564;
Best Local Similarity 59.2%; Pred. No. 4.2e-60;
Matches 539; Conservative 0; Mismatches 359; Indels 12; Gaps 2;
QY 952 GATTAACGATTAAGTGGCTGGCATGGGCAACCAAGAGCTGGGAATCTTGACAAAGTAT 1011
DB 288 GATGCTCAGTACTAGTGTAGGCAAGGCAACCAAGAACTTCTGAGTATTTCAAGTATAT 347
QY 1012 GAGGCTCTTAGAGCAGCCGCTTCTATGCTCAACAGGGCCATCTGGAGTGAAGTCTTG 1071
DB 348 GCTGCGACCAAGACAGCCATGATATGTCAGGCGACCGTGGCATGAGTGTGCTA 407
QY 1072 ATGTTTGAAGCAGTGGCTGCTATTTGAGGCGCAACGCTCCACCGGAGTAGCT 1131
DB 408 ATATTTAAAGCTCTGCTGTGGCTATATGAGGCAAGACGTTCTATTAAGCATTTTAT 467
QY 1132 GAGATGGGGTTAGATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1188
DB 468 GATCAAAAGAACAGCAGGAGACCTTGCAAGAAATGCAAGGCTCTTTCTTAACTGGTGG 527
QY 1189 GTTGGCAACTGTATGCTTCTTCAAGCAAGACAGATTTGACATATTTCAATCAAC 1248
DB 528 AAGAGCAATATATCGGTTCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 587
QY 1249 TCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1308
DB 588 TGCCAAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 647
QY 1309 GAGCTGAGGCAAGCTCTTGAAGCAATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTCTCA 1368

DB 648 CAGATGAACGAATATGCGAGATATATCAACACTATTTCTGAAGAACAGAGTGT 707
QY 1369 AAACAGAACAGCAGCCAGAGTGTCTTGAAGATTTCTGAAATTTATGAGGAGAGAGCTG 1428
DB 708 AAGACAGACAGAGGCTCTTAAGTTGTAGAGAAACCTGGGTATTAACCGAACTT 767
QY 1429 CGTAAGCTGACAGAGATTAATCGGATCTGAGACAGAGAACTAAGATCAGCATGAC 1488
DB 768 CGAGAGCTATGGAAGAGATATTTCTGAGAGTAAAGCTTAAGAGAGAGCACTGAG 827
QY 1489 AACGGAAGAGATGATGATCAGACAGAGCTTTTTCATGATTCATCAATCAACAGTTCAT 1548
DB 828 TAGAGAGAGAGATGAAATCCAGAGAAATCTTCATGATCAGATGAGAACTTC 887
QY 1549 GAAAGAGAGACCCAGAGAGAGAGATTTTTCAGATTTTTCAGACAGAGAGAGAGCTGCAAG 1608
DB 888 AAGGCCACAG 947
QY 1609 GTTGTGGCCAGCAGCAGCAAGCAATTAATCCCTTGAAGATGACGATTCGGAAGAGA 1668
DB 948 GCT-----AGACGTGTGATGATGATTTCTGGGACTAATGAAATGCGAGGCTAAG 998
QY 1669 GCTGAGAGAGTGTCAAGCTTCATGAGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1728
DB 999 AAGGAGAGGTAACGCGTTCATAGAGTGCAGAGTTAAGAGAGAGAGAGAGAGAGAG 1058
QY 1729 GAGAGAGAGATGCTGATTAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1788
DB 1059 GAAAG 1118
QY 1789 CACGAGAGATATTTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1848
DB 1119 ATGCAAG 1178
QY 1849 AAGCATGGCC 1858
DB 1179 AAGCATGAGC 1188

RESULT 12
AK061975 1214 bp mRNA linear PLN 24-JUL-2003
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:001-043-A02, full
DEFINITION
ACCESSION AK061975
VERSION AK061975.1 GI:32971993
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzaceae; Oryza.
1
The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team;
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikura, J.,
Ikedo, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Aizawa, K., Arakawa, T., Fukuda, S.,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Ishii, Y., Itoh, M.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Ota, Y.,
Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Otsu, N., Shitaki, T.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)
 MEDLINE 22752273
 PUBMED 12869764
 REFERENCE
 AUTHORS 2 (bases 1 to 1214)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imocani,K., Iribiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kangawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurotaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Nikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osafo,N., Ota,Y., Otsu,Y., Ryo,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shikazume,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
 Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
 This clone is one of the 28k full-length cDNA clones from japonica rice.
 URL: http://cdna01.dna.affrc.go.jp/cDNA/
 NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
 PAIS Genome Sequencing & Analysis Group: Otsu,Y., Iida,Y., Fujimura,T., Ikeda,R., Iribiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurotaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Nikura,J., Oka,M., Ryo,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
 Genomic Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Iida,Y., Imamura,K., Imocani,K., Iribiki,J., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osafo,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasakawa,D., Sato,K., Satoh,K., Shibata,K., Shikazume,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.
 Location/Qualifiers
 1. 1214
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiVar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="001-043-A02"

ORIGIN

Query Match 15.4%; Score 288.8; DB 8; Length 1214;
 Best Local Similarity 58.7%; Prem. No. 2e-59;
 Matches 544; Conservative 0; Mismatches 367; Indels 15; Gaps 2;
 QY 934 ACTAGACTGATTAAGACGATTAAGTGTGCTGGATGGCAACGAAGCTGCTG 993

FEATURES

source
 1. 1214
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiVar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="001-043-A02"

Db 2 ACTGATTTGGAACAAGATGAGATGATTAAGTGGAAAGGATATGGGGAACCAAGAGCTAAAT 61
 QY 994 GAATCTTCGACAAAGATGAGATGCTTTAGAGACCGCATTCATATGCTCCACAGGCCCAT 1053
 Db 62 GATTATTATTAAGTAATATGCTCCAGCAAGACGCAAGCCTATATGCTCCAAATGGGCAC 121
 QY 1054 CTTGGATGAGTGTCTGATGTTTGGAGACAGTGCACATGCTATTTGGAGGCCCAAGC 1113
 Db 122 CTTGGATGAGTGTCTGATGTTTGGAGACAGTGCACATGCTATTTGGAGGCCCAAGC 181
 QY 1114 CTCCACCGGAGTGTGAGATGAGTGGGTTAGATTAATTCCTG-----GCTCAGAA 1167
 Db 182 TTGATGATGATCACTTTGTTGCTGCAAGAAACAGACAGGAATATATGAAACAGTGTCTCAAG 241
 QY 1168 CGCAGATGTTTTCGAGAGTGTGGCCCACTGTATGCTCTCTTCCACAGCAAGAT 1227
 Db 242 GTTACATTTCTTACCTGTGTGGAAAGGACGCTGTATGTTCTTACCTTCAAAAGATGAC 301
 QY 1228 CTGACATATTCATCAACTCTCAAGGCAAAACAAAGCTGAATTCAGTTGAATCA 1287
 Db 302 ATGAGACATTTCAATAGACATTTGCCAGCAAAATCCCTTTGAATATGATGATCT 361
 QY 1288 TACCAAGATGCTGTGTAAAGAGCTGAGCAGATCTTGAGACAAATGAGCAGCTGAAC 1347
 Db 362 TATTAATGATGATGTGTATCCCAATGAAAGATGAGGAAACCAACAGCAATGAAAT 421
 QY 1348 TACTTTAACAACAGCTCCAAACGAACCAAGCCCAAGTCTTGAAGAAATCTCG 1407
 Db 422 TACCTGAACAAACAAATGTTTAAAGGACAGCACTTAAATTTAGTAAAGATTAATTTG 481
 QY 1408 GAAATTAATGAGGAGAGAGCTGGGTAGAACTGACAGAGATTAATCGATCTGTAGACAGAG 1467
 Db 482 AGTGTGTTTATCCCAAGAACTTTAGGAGACAAATGGAAGAAATCTATTGTCGAATATA 541
 QY 1468 ACTAAGATGACAGATGAACAGAAACAGGAGAGATGATGACACAGCAGGTTTTTCATG 1527
 Db 542 GCCAAAGAAACCACTTGAGATATGAGAAAGATGAATGATCAACAGAGAAATTTTCCAT 601
 QY 1528 GATTCATTAACACAGATTCATGAAGAGACGCAAGAGAGAAATTTCCAGATGTT 1587
 Db 602 GATCAATGAAGAAATTTACAAAGGCCACAGAGAGAAAGAAATCAAGTTTGAAAGTTG 661
 QY 1588 CAGCAGCAGAAACGTCGCAAGGTTGTGGCCAGCAGCAGACAGAACTTAATCCCTTAG 1647
 Db 662 CTGCAAGAGAGCGCAAGGCT-----AGACAAAGTATTTGATCTTGAAGT 712
 QY 1648 AATGACGATTTGCCAAGAGAGCTGAGAGTGTCAAGCTTCATCGATTTCAAGAGAA 1707
 Db 713 ACTGAAGATCGCAGGCAAGGAAAGAAAGAAATTCAGAACTTCATGACTGCCAGTGAAG 772
 QY 1708 GAGATGAGAGAGTTTGTGAAGAGAGGAGATCTGATTAAGATTAAGAGAAAGATG 1767
 Db 773 GATGTTGAGAAATTTGACACCCCAACGAGCAAACTGATCAAAATTTGATGAAGAGAAAG 832
 QY 1768 GAAGCATGAAGAAGGATCAGCAGAGATTTTGTCTGAGAGAAAGATTTGATGAG 1827
 Db 833 GTGAAGCTCAAGAGAGATTAAGTCTGGCAAGGAATTTGACTTGAAGAGAGCTGGACACA 892
 QY 1828 GCTTTGGAACAGCTCATGTACAGCA 1853
 Db 893 GCCCTCATCTCCCTGATGATGAATAACA 918
 RESULT 13
 AF469493 513 bp mRNA linear PLN 02-JUL-2003
 LOCUS Triticum aestivum H07R mRNA, partial cds.
 DEFINITION AF469493
 ACCESSION AF469493.1 GI:3240075
 VERSION
 KEYWORDS
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE
AUTHORS
JOURNAL
TITLE
AUTHORS
JOURNAL
TITLE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 513)
Li,J.R., Wang,F., Li,O.Z. and Zhang,X.S.
Gene isolation and expression of a new Zn-finger
Unpublished
2 (bases 1 to 513)
Li,J.R., Wang,F., Li,O.Z. and Zhang,X.S.
Direct Submission
Submitted (18-JAN-2002) College of Life Science, Shandong
Agricultural University, Dai zong Street 61, Tatan, Shandong

FEATURES	Location/Qualifiers
source	1. . 513
	/organism="Triticum aestivum"

CDS
 <1. .513
 /note="similar to HOPR of Hordeum vulgare"

```
<1.513
/note="similar to HtRr of Hordeum vulgare"
/codon_start=1
/product="HtRr"
/protein_id="AAP060.1"
/translation="WLDWNGDLELPSDYATKTHAYGPGGHHGMSVILPESFSAVG
YMEARLHKHETIDRTRIDWNRKRPFLPGSGPOLYGLAKEMETFNHCOGKSKS
LKYMRSGINEVVAQKMSEDNOQLATLNKKVTKYKEDSKVVEETLVATITQKLETV
ENI FVRSKA"
```

ORIGIN

[illegible]

RESULT 14	CNS08C0/c	CNS08C0	LOCUS	DEFINITION
	91053 bp	DNA	linear	FIN 21-NOV-2003
				Oryza sativa chromosome 12 of library Monsanto
				Oryza sativa chromosome 12 of cultivar Nippondare of ssp. japonica of Oryza
				from chromosome 12 of cultivar Nippondare of ssp. japonica of Oryza
				sativa (rice), complete sequence.

ACCESSION	AL83J1811
VERSION	AL83J1811.4
KEYWORDS	GI:28892661
SOURCE	HTG.
ORGANISM	Oryza sativa (japonica cultivar-group)
	Oryza sativa (japonica cultivar-group)
	Oryza sativa (japonica cultivar-group)
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Eudicotyledons; Magnoliopsida; Poales; Poaceae;
	Monoclony; Monoclonal; Monoclonal; Monoclonal; Monoclonal; Monoclonal;

COMMENT
On Mar 9, 2003 this sequence version appeared in
Center: Genoscope / Centre National de Sequencage

```

The following sequence is oriented from the T7 to the SP6 end. The
nucleotide sequence of this BAC clone was generated by combining
Monsanto, Syngenta and Genosco BAC sequencing data.
Upstream BAC (overlapping the T7 end) : OSUNBA0029N15 (AC=BX536367)
Downstream BAC (overlapping the SP6 end) : OJ1102_B11 (AC=AL7113801)
----- Finishing boundaries -----
FINISHED SEGMENT STARTS AT BASE 1
FINISHED SEGMENT ENDS AT BASE 75065
-----

```

```

FEATURES
    source
        location/Qualifiers
            1..91053
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="genomic DNA"
            /cultivar="Nipponbare"
            /sub_species="japonica"
            /db_xref="taxon:39947"
            /chromosome="12"
            /clone="OJ1561.A05"
            /clone_1fb="Monsanto"

```

Query Match	9.0%	Score 169.4;	DB 8;	Length 91053;
Best Local Similarity	61.1%;	Pred. No. 2.6e-30;		
Matches 310;	Conservative	0;	Mismatches 191;	Indels 6; Gaps 2

Qy	460	TTTCCTGACGTGGAGAGATGATGTGGTAATGCTTCGAGAGAGAGATTCCTCATCT	519
Db	81521	TTGTCTGGTTGTGATCCTGAAACAGACATGCTGMAAGTTTATGACATCGACGATGAT	81462
Qy	520	TTGGATGATTTCTGATGACGACCTTGCAAGTATGATTTATGACTCGATGTGTGACAAAAG	579
Db	81461	AATGATGATGATATATAGTACGATTTGAGTGATGACTATATCTTATCATCAATGAGAAA	81402
Qy	580	AGCCATGATTCACGAAAGCAGATTAAGTGTTCAAAAAGTCTTTTGACAGCTTGATAGC	639
Db	81401	AGTTTGAAGCTTCGAAAAATCACAAGTTGTTCAAAGTTTCTTCGAAAGCTCGAGGCG	81342
Qy	640	TTTGATCATGAGACAGATTAATGAACCAAGACAGGACATGGCATTTGTCCAGTTGTTCGAAC	699
Db	81341	TTTAAGTGTGAACAGTTAATGAACCAACTGAGCAATGGCATTTGCCAGATTCAGAAAAAT	81288

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 14:24:13 ; Search time 494.227 Seconds
(without alignments)
16142.601 Million cell updates/sec

Title: US-10-030-829-2

Perfect score: 1878
Sequence: 1 atgagcttcaggcgcgtcc.....ttcacatgaagatgatga 1878

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1878	100.0	1878	4	AAFP25374
2	1462	77.8	3275	4	AAFP25373
3	75	4.0	272	7	ABX81661
4	74.4	4.0	274	7	ABX82003
5	63.6	3.4	310	6	ABLT1922
6	58.8	3.1	3489	3	AAAT0290
7	58.8	3.1	3489	4	AAFP2901
8	58.8	3.1	3489	6	ABX83487
9	58.8	3.1	32207	2	AAV73805
10	58.8	3.1	137507	2	AAV19941
11	56.2	3.0	1080	6	ABT199537
12	55	2.9	1686	2	AAO87587
13	55	2.9	4000	2	AAAT91902
14	53.2	2.8	1140	3	AAAC40071
15	53.2	2.8	16442	2	AAAT83006
16	53	2.8	3399	2	AAAT05868
17	52.6	2.8	700	4	AAAT20076
18	52.6	2.8	700	4	ABAT65105
19	52.6	2.8	700	4	ABAT45274
20	52.6	2.8	700	4	ABAT47220
21	52.6	2.8	700	4	ABAT32209
22	52.6	2.8	700	4	AAK39264
23	52.6	2.8	700	4	AAK13527

24	52.6	2.8	700	4	ABAT38852	ABAT38852 Human liv
25	52.6	2.8	700	5	AAAT05782	AAAT05782 Probe #57
26	52.6	2.8	700	6	AAAT33350	AAAT33350 Human.gen
27	52.2	2.8	2000	7	ADAT19318	ADAT19318 Rice gene
28	52	2.8	766	4	AAI22114	AAI22114 Probe #12
29	52	2.8	766	4	ABAT67193	ABAT67193 Human foe
30	52	2.8	766	4	AAAT47409	AAAT47409 Probe #16
31	52	2.8	766	4	ABAT49279	ABAT49279 Human bre
32	52	2.8	766	4	ABAT44288	ABAT44288 Probe #12
33	52	2.8	766	5	AAAT07812	AAAT07812 Probe #78
34	52	2.8	1944	4	AAI12874	AAI12874 Probe #28
35	52	2.8	1944	4	ABAT54575	ABAT54575 Human foe
36	52	2.8	1944	4	AAAT34231	AAAT34231 Probe #29
37	52	2.8	1944	4	ABAT44123	ABAT44123 Human bre
38	52	2.8	1944	4	ABAT24359	ABAT24359 Probe #28
39	52	2.8	1944	5	AAAT02791	AAAT02791 Probe #27
40	51.6	2.7	2093	9	ADAT63512	ADAT63512 Human CDN
41	51.2	2.7	693	5	AAAT74240	AAAT74240 DNA encod
42	51.2	2.7	693	5	AAAT90715	AAAT90715 DNA encod
43	51	2.7	267	4	AAAT45604	AAAT45604 Human bon
44	51	2.7	267	4	AAAT19599	AAAT19599 Human bra
45	51	2.7	267	4	ABAT5294	ABAT5294 Human liv

ALIGNMENTS

RESULT 1
AAFP25374
ID AAFP25374 standard; cDNA; 1878 BP.
XX
AC AAFP25374;
XX
DT 15-MAY-2001 (first entry)
XX
DE Nucleotide sequence of the Arabidopsis SGS3 polypeptide.
XX
KW SGS3 gene; post-transcriptional inactivation; RNA degradation;
XX viral resistance; resistance; fatty acid content; protein content; sb.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 1..1878 /tag= a
FT FT /product= "SGS3"
XX
PN WO200105951-A2.
XX
PD 25-JAN-2001.
XX
PF 13-JUL-2000; 2000MO-FR002052.
XX
PR 16-JUL-1999; 99FR-00009417.
XX 26-JAN-2000; 2000FR-00001006.
XX
PA (AVET) AVENTIS CROSCIENCE SA.
XX (INRG) INST NAT RECH AGRONOMICQUE.
XX
PI Beclin C, Elmeyan T, Vaucheret H;
XX WPI; 2001-159529/16.
XX DR P-FSDB; AAB31798.
XX
XX New SGS3 gene from Arabidopsis thaliana, useful for increasing virus
XX resistance in plants and, when inhibited, for increasing transgene
XX expression.
XX
XX Claim 1; Page 32-35; 36pp; French.
XX
XX The present sequence encodes the Arabidopsis thaliana SGS3 polypeptide.
XX The SGS3 gene is essential for post-transcriptional inactivation
XX (degradation of RNA) and for resistance to viruses. Overexpression of

CC SSS3 results in plants with increased resistance to viruses, while
 CC inactivation of SSS3 in transgenic plants (e.g. by expressing the level of
 CC RNA, by mutation or by homologous recombination) increases the level of
 CC the transgene product. This product may e.g. impart resistance (to
 CC herbicide, insects or pathogens), alter contents of essential fatty acids
 CC or proteins, or is pharmaceutically active, e.g. an immunoglobulin or
 CC interferon

XX Sequence 1878 BP; 575 A; 332 C; 554 G; 417 T; 0 U; 0 Other;

Query Match 100.0%; Score 1878; DB 4; Length 1878;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1878; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTTCTAGGCTGCTGCTCAATGCTTAAGAAAACAGTTCAGGGGTGTTATAGGCT 60
 Db 1 ATGAGTTCTAGGCTGCTGCTCAATGCTTAAGAAAACAGTTCAGGGGTGTTATAGGCT 60
 QY 61 GAGGTGAACAGTGTGTTCAAGTTTGCGAGGACGAGCTGGCTTCTTCAAGATGAT 120
 Db 61 GAGGTGAACAGTGTGTTCAAGTTTGCGAGGACGAGCTGGCTTCTTCAAGATGAT 120
 QY 121 GAGAGAGAGTGGAGGTCAATTTCCAGAAACAAAGAAACAGAGAAACATTTCTGGA 180
 Db 121 GAGAGAGAGTGGAGGTCAATTTCCAGAAACAAAGAAACAGAGAAACATTTCTGGA 180
 QY 121 GAGAGAGAGTGGAGGTCAATTTCCAGAAACAAAGAAACAGAGAAACATTTCTGGA 180
 Db 121 GAGAGAGAGTGGAGGTCAATTTCCAGAAACAAAGAAACAGAGAAACATTTCTGGA 180
 QY 181 AAAAATCTGGGTTCTCAGAAATTCCTCTAGAGCTTGGGGTGGTCAAGCAAGG 240
 Db 181 AAAAATCTGGGTTCTCAGAAATTCCTCTAGAGCTTGGGGTGGTCAAGCAAGG 240
 QY 181 AAAAATCTGGGTTCTCAGAAATTCCTCTAGAGCTTGGGGTGGTCAAGCAAGG 240
 Db 181 AAAAATCTGGGTTCTCAGAAATTCCTCTAGAGCTTGGGGTGGTCAAGCAAGG 240
 QY 241 AGAGGTACCAACGTAATCTGGAGAGAAACATGATTCGGAGAGAGTAACGGCAATGT 300
 Db 241 AGAGGTACCAACGTAATCTGGAGAGAAACATGATTCGGAGAGAGTAACGGCAATGT 300
 QY 241 AGAGGTACCAACGTAATCTGGAGAGAAACATGATTCGGAGAGAGTAACGGCAATGT 300
 Db 241 AGAGGTACCAACGTAATCTGGAGAGAAACATGATTCGGAGAGAGTAACGGCAATGT 300
 QY 301 CGGGGCAATTCAGCTTAACATATCTGGTGGGAGCAGCGCTTGACAGAAATGATATAC 360
 Db 301 CGGGGCAATTCAGCTTAACATATCTGGTGGGAGCAGCGCTTGACAGAAATGATATAC 360
 QY 301 CGGGGCAATTCAGCTTAACATATCTGGTGGGAGCAGCGCTTGACAGAAATGATATAC 360
 Db 301 CGGGGCAATTCAGCTTAACATATCTGGTGGGAGCAGCGCTTGACAGAAATGATATAC 360
 QY 361 AACTTTGTGGACCCCACTGTATCTGCCCTCTTTGGAAGAGATGAAATTTGGCAG 420
 Db 361 AACTTTGTGGACCCCACTGTATCTGCCCTCTTTGGAAGAGATGAAATTTGGCAG 420
 QY 361 AACTTTGTGGACCCCACTGTATCTGCCCTCTTTGGAAGAGATGAAATTTGGCAG 420
 Db 361 AACTTTGTGGACCCCACTGTATCTGCCCTCTTTGGAAGAGATGAAATTTGGCAG 420
 QY 421 GCAAGAGAGAGTTCCTGCTCAAGACACAGCTGCGAGAGTTTCTGAGTGAAGATAT 480
 Db 421 GCAAGAGAGAGTTCCTGCTCAAGACACAGCTGCGAGAGTTTCTGAGTGAAGATAT 480
 QY 421 GCAAGAGAGAGTTCCTGCTCAAGACACAGCTGCGAGAGTTTCTGAGTGAAGATAT 480
 Db 421 GCAAGAGAGAGTTCCTGCTCAAGACACAGCTGCGAGAGTTTCTGAGTGAAGATAT 480
 QY 481 GTGGATTAATGCTCTGAGAGAGAGATGATTCGATCTTTGATGATTCGATGACAC 540
 Db 481 GTGGATTAATGCTCTGAGAGAGAGATGATTCGATCTTTGATGATTCGATGACAC 540
 QY 481 GTGGATTAATGCTCTGAGAGAGAGATGATTCGATCTTTGATGATTCGATGACAC 540
 Db 481 GTGGATTAATGCTCTGAGAGAGAGATGATTCGATCTTTGATGATTCGATGACAC 540
 QY 541 CTTCGAGATGATTAATGATCTGATGATGATGATGATGATGATGATGATGATGATG 600
 Db 541 CTTCGAGATGATTAATGATCTGATGATGATGATGATGATGATGATGATGATGATG 600
 QY 541 CTTCGAGATGATTAATGATCTGATGATGATGATGATGATGATGATGATGATGATG 600
 Db 541 CTTCGAGATGATTAATGATCTGATGATGATGATGATGATGATGATGATGATGATG 600
 QY 601 AATTAAGGTGTTCAAAAAGTTCTTTGGCAGCTTGATGATGATGATGATGATGAT 660
 Db 601 AATTAAGGTGTTCAAAAAGTTCTTTGGCAGCTTGATGATGATGATGATGATGAT 660
 QY 601 AATTAAGGTGTTCAAAAAGTTCTTTGGCAGCTTGATGATGATGATGATGATGAT 660
 Db 601 AATTAAGGTGTTCAAAAAGTTCTTTGGCAGCTTGATGATGATGATGATGATGAT 660
 QY 661 GAACACACAGAGAGAGTGGCATTTGTCAGCTTTGTCAGAAACAGTGGTGCATGATG 720
 Db 661 GAACACACAGAGAGAGTGGCATTTGTCAGCTTTGTCAGAAACAGTGGTGCATGATG 720
 QY 661 GAACACACAGAGAGAGTGGCATTTGTCAGCTTTGTCAGAAACAGTGGTGCATGATG 720
 Db 661 GAACACACAGAGAGAGTGGCATTTGTCAGCTTTGTCAGAAACAGTGGTGCATGATG 720
 QY 721 TATTAACCTGACCCCTCTAATGCTCATGCGAGCAAAAAGAGCTAGGCAATTAAGCTC 780
 Db 721 TATTAACCTGACCCCTCTAATGCTCATGCGAGCAAAAAGAGCTAGGCAATTAAGCTC 780
 QY 721 TATTAACCTGACCCCTCTAATGCTCATGCGAGCAAAAAGAGCTAGGCAATTAAGCTC 780
 Db 721 TATTAACCTGACCCCTCTAATGCTCATGCGAGCAAAAAGAGCTAGGCAATTAAGCTC 780
 QY 781 CATAGAGAAATGCTGAAGTTTGAAGAAAGATCTAAGATGAGAGGCGCATGTCATT 840
 Db 781 CATAGAGAAATGCTGAAGTTTGAAGAAAGATCTAAGATGAGAGGCGCATGTCATT 840
 QY 781 CATAGAGAAATGCTGAAGTTTGAAGAAAGATCTAAGATGAGAGGCGCATGTCATT 840
 Db 781 CATAGAGAAATGCTGAAGTTTGAAGAAAGATCTAAGATGAGAGGCGCATGTCATT 840
 QY 841 CCTTGTGTGATGATTTAATGAGAGAGGATTTGGGTGAGATGAAAGATTAATGAA 900
 Db 841 CCTTGTGTGATGATTTAATGAGAGAGGATTTGGGTGAGATGAAAGATTAATGAA 900

QY 901 ATTGTGCTCTCCATGCTCATCATGTAATCTAGCTGATTAAGACATTAACAT 960
 Db 901 ATTGTGCTCTCCATGCTCATCATGTAATCTAGCTGATTAAGACATTAACAT 960
 QY 961 AATGTGCTGCTCCATGCTCATCATGTAATCTAGCTGATTAAGACATTAACAT 960
 Db 961 AATGTGCTGCTCCATGCTCATCATGTAATCTAGCTGATTAAGACATTAACAT 960
 QY 961 AATGTGCTGCTCCATGCTCATCATGTAATCTAGCTGATTAAGACATTAACAT 960
 Db 961 AATGTGCTGCTCCATGCTCATCATGTAATCTAGCTGATTAAGACATTAACAT 960
 QY 1021 AGAGACGCCATTCCTATGCTTCAAGGCGCATGCTGGATGAGTCTGATGTTGAG 1080
 Db 1021 AGAGACGCCATTCCTATGCTTCAAGGCGCATGCTGGATGAGTCTGATGTTGAG 1080
 QY 1021 AGAGACGCCATTCCTATGCTTCAAGGCGCATGCTGGATGAGTCTGATGTTGAG 1080
 Db 1021 AGAGACGCCATTCCTATGCTTCAAGGCGCATGCTGGATGAGTCTGATGTTGAG 1080
 QY 1081 AGAGAGCCATGCTGATTTGAGAGCGCAACCGCTTCAACCGGAGTTAGCTGATGAG 1140
 Db 1081 AGAGAGCCATGCTGATTTGAGAGCGCAACCGCTTCAACCGGAGTTAGCTGATGAG 1140
 QY 1081 AGAGAGCCATGCTGATTTGAGAGCGCAACCGCTTCAACCGGAGTTAGCTGATGAG 1140
 Db 1081 AGAGAGCCATGCTGATTTGAGAGCGCAACCGCTTCAACCGGAGTTAGCTGATGAG 1140
 QY 1141 TTGATTAAGAAATGCTGCTGGGTCAAGAGCGCATGATCTTTCTGAGAGTTCGCAACTG 1200
 Db 1141 TTGATTAAGAAATGCTGCTGGGTCAAGAGCGCATGATCTTTCTGAGAGTTCGCAACTG 1200
 QY 1141 TTGATTAAGAAATGCTGCTGGGTCAAGAGCGCATGATCTTTCTGAGAGTTCGCAACTG 1200
 Db 1141 TTGATTAAGAAATGCTGCTGGGTCAAGAGCGCATGATCTTTCTGAGAGTTCGCAACTG 1200
 QY 1201 TATGCTTCTTGGACAGAGCAAGATCTGACATATTCATCAACCTCTCAAGCAAA 1260
 Db 1201 TATGCTTCTTGGACAGAGCAAGATCTGACATATTCATCAACCTCTCAAGCAAA 1260
 QY 1201 TATGCTTCTTGGACAGAGCAAGATCTGACATATTCATCAACCTCTCAAGCAAA 1260
 Db 1201 TATGCTTCTTGGACAGAGCAAGATCTGACATATTCATCAACCTCTCAAGCAAA 1260
 QY 1261 ACAAGGCTGAATTCAGTTGAATCATACAGAGATGTTGTAAGAGCTGAGGAG 1320
 Db 1261 ACAAGGCTGAATTCAGTTGAATCATACAGAGATGTTGTAAGAGCTGAGGAG 1320
 QY 1261 ACAAGGCTGAATTCAGTTGAATCATACAGAGATGTTGTAAGAGCTGAGGAG 1320
 Db 1261 ACAAGGCTGAATTCAGTTGAATCATACAGAGATGTTGTAAGAGCTGAGGAG 1320
 QY 1321 ATCTGAGAGCAATCAGAGCTGATCTAATTAAGAACAGCTTCAAAACAGAACAG 1380
 Db 1321 ATCTGAGAGCAATCAGAGCTGATCTAATTAAGAACAGCTTCAAAACAGAACAG 1380
 QY 1321 ATCTGAGAGCAATCAGAGCTGATCTAATTAAGAACAGCTTCAAAACAGAACAG 1380
 Db 1321 ATCTGAGAGCAATCAGAGCTGATCTAATTAAGAACAGCTTCAAAACAGAACAG 1380
 QY 1381 CACGCCAAGTGTCTTGAAGAAATCTTGAATTAAGAGCAAGCTGCTGAAGCTGCA 1440
 Db 1381 CACGCCAAGTGTCTTGAAGAAATCTTGAATTAAGAGCAAGCTGCTGAAGCTGCA 1440
 QY 1381 CACGCCAAGTGTCTTGAAGAAATCTTGAATTAAGAGCAAGCTGCTGAAGCTGCA 1440
 Db 1381 CACGCCAAGTGTCTTGAAGAAATCTTGAATTAAGAGCAAGCTGCTGAAGCTGCA 1440
 QY 1441 GAGATTAATCGATGCTGAGACAGAGAACTAAGATGAGATGAGAAACAGGGAAGAG 1500
 Db 1441 GAGATTAATCGATGCTGAGACAGAGAACTAAGATGAGATGAGAAACAGGGAAGAG 1500
 QY 1441 GAGATTAATCGATGCTGAGACAGAGAACTAAGATGAGATGAGAAACAGGGAAGAG 1500
 Db 1441 GAGATTAATCGATGCTGAGACAGAGAACTAAGATGAGATGAGAAACAGGGAAGAG 1500
 QY 1501 ATGGATGCAACAGAGGTTTTCATGATTAATCAACAGATCTTGAAGAGAGAGAC 1560
 Db 1501 ATGGATGCAACAGAGGTTTTCATGATTAATCAACAGATCTTGAAGAGAGAGAC 1560
 QY 1501 ATGGATGCAACAGAGGTTTTCATGATTAATCAACAGATCTTGAAGAGAGAGAC 1560
 Db 1501 ATGGATGCAACAGAGGTTTTCATGATTAATCAACAGATCTTGAAGAGAGAGAC 1560
 QY 1561 GCAAGAGAGAGAAATTTGAGATGTTGAGACAGAGAACTGCAAGCTTGTGGCAG 1620
 Db 1561 GCAAGAGAGAGAAATTTGAGATGTTGAGACAGAGAACTGCAAGCTTGTGGCAG 1620
 QY 1561 GCAAGAGAGAGAAATTTGAGATGTTGAGACAGAGAACTGCAAGCTTGTGGCAG 1620
 Db 1561 GCAAGAGAGAGAAATTTGAGATGTTGAGACAGAGAACTGCAAGCTTGTGGCAG 1620
 QY 1621 CAGCAGCAGAACTTAATCCCTCTAGCAATGACATTCGCAAGAGAGCTGAGAGAGT 1680
 Db 1621 CAGCAGCAGAACTTAATCCCTCTAGCAATGACATTCGCAAGAGAGCTGAGAGAGT 1680
 QY 1621 CAGCAGCAGAACTTAATCCCTCTAGCAATGACATTCGCAAGAGAGCTGAGAGAGT 1680
 Db 1621 CAGCAGCAGAACTTAATCCCTCTAGCAATGACATTCGCAAGAGAGCTGAGAGAGT 1680
 QY 1681 TCAAGCTTCATGAGTTTCAAGAGAAAGAGAGAGTTTGTGAAGAGAGAGAGT 1740
 Db 1681 TCAAGCTTCATGAGTTTCAAGAGAAAGAGAGAGTTTGTGAAGAGAGAGAGT 1740
 QY 1681 TCAAGCTTCATGAGTTTCAAGAGAAAGAGAGAGTTTGTGAAGAGAGAGAGT 1740
 Db 1681 TCAAGCTTCATGAGTTTCAAGAGAAAGAGAGAGTTTGTGAAGAGAGAGAGT 1740
 QY 1741 CTGATTAAGAAATCAAGAGAAAGATGAGAGATGAGAGAGAGAGAGATCAAGAGAGATA 1800
 Db 1741 CTGATTAAGAAATCAAGAGAAAGATGAGAGATGAGAGAGAGAGAGATCAAGAGAGATA 1800
 QY 1741 CTGATTAAGAAATCAAGAGAAAGATGAGAGATGAGAGAGAGAGAGATCAAGAGAGATA 1800
 Db 1741 CTGATTAAGAAATCAAGAGAAAGATGAGAGATGAGAGAGAGAGAGATCAAGAGAGATA 1800
 QY 1801 TTTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAACAGCTCATGTAAGAGATGAGCTT 1860
 Db 1801 TTTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAACAGCTCATGTAAGAGATGAGCTT 1860
 QY 1801 TTTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAACAGCTCATGTAAGAGATGAGCTT 1860
 Db 1801 TTTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAACAGCTCATGTAAGAGATGAGCTT 1860
 QY 1861 CACAATGAAGATGATGA 1878
 Db 1861 CACAATGAAGATGATGA 1878

RESULT 2
 AAF25373
 ID AAF25373 standard; DNA; 3275 BP.


```

QY 1068 TCTGATGTTGAGAGAGAGTGCACATGCTATTTGAGGCCGCAACGCTCCACCGGAGTT 1127
DB 1836 TCTGATGTTGAGAGAGAGTGCACATGCTATTTGAGGCCGCAACGCTCCACCGGAGTT 1895
QY 1128 AGCTGATGAGGGGTTAGATAGAAATGCTGGGGTCAAGAACGCGATATGTTTCTGAGG 1187
DB 1896 AGCTGATGAGGGGTTAGATAGAAATGCTGGGGTCAAGAACGCGATATGTTTCTGAGG 1955
QY 1188 TGTTCGGCCAACTGTATGAGCTTCTTCCAGAGAGCAAGATCTGGAATATTCATCAACA 1247
DB 1956 TGTTCGGCCAACTGTATGAGCTTCTTCCAGAGAGCAAGATCTGGAATATTCATCAACA 2015
QY 1248 CTCTCA----- 1253
DB 2016 CTCTCAAGTTCTCTCCCAAGAAATTTGATATATGCTTTTATGTTTGTCAATTGGAAT 2075
QY 1254 -----AGG 1256
DB 2076 TTAAGTTTGTGTGTCGGTAAATGATCATCTGTATATATCTATGATTCATTAGG 2135
QY 1257 CAAAACAGGCTGAAATTTGAGTTGAAATCATACCAAGATGTTTAAAGAGCTGAG 1316
DB 2136 CAAAACAGGCTGAAATTTGAGTTGAAATCATACCAAGATGTTTAAAGAGCTGAG 2195
QY 1317 GCAGATCTCTGAGAGCAATCAGAGCTGAATCTTTAAGAACAACTCTCAAAAACAGAA 1376
DB 2196 GCAGATCTCTGAGAGCAATCAGAGCTGAATCTTTAAGAACAACTCTCAAAAACAGAA 2255
QY 1377 CAAAGCAGCAGAGTCTTGAAGAAATCTGTGAAATTTATGAGGAGAGCTGGTGAAC 1436
DB 2256 CAAAGCAGCAGAGTCTTGAAGAAATCTGTGAAATTTATGAGGAGAGAGCTGGTGAAC 2315
QY 1437 TCCAGAGCAATATCCGATCGTGAAGAGAACTAATAGCCAGCAAGAAACAGAGGA 1496
DB 2316 TCCAGAGCAATATCCGATCGTGAAGAGAACTAATAGCCAGCAAGAAACAGAGGA 2375
QY 1497 AG----- 1498
DB 2376 AAGAGTATGATTTTCTTGAAGAAATCAAACTTGAATTTGTATTAATCTACTGATTC 2435
QY 1499 -----AGATGATGACACGA 1514
DB 2436 CATTTTGAATATATGTCCAAAAAAACTGTGTGTTGTTGAAGATGATGATCAACGA 2495
QY 1515 CAGGTTTTCATGATTCATCAATCAAGATCCATGAAAGAGAGAGCAAGAGAGAGAA 1574
DB 2496 CAGGTTTTCATGATTCATCAATCAAGATCCATGAAAGAGAGAGAGAGAGAGAGAA 2555
QY 1575 TTTTCGATGTTGACAGAGAGAGAGAGTGTGTTGGCCAGACAGAGAGAAAT 1634
DB 2556 TTTTCGATGTTGACAGAGAGAGAGAGTGTGTTGGCCAGACAGAGAGAAAT 2615
QY 1635 TAATCCCTTAGCAATGACGATGCGCAAG----- 1665
DB 2616 TAATCCCTTAGCAATGACGATGCGCAAGAGGTATATGTAATAATTAATCCCT 2675
QY 1666 ----- 1665
DB 2676 CTGGCGTTTTTTTCAAACTTAAGATTAATTCGGTTTTGATTCCTTTCG 2735
QY 1666 -AGAGCTGAGAGAGTCAAGCTTCAATGAGTTTCAAGAGAAAGAGATGAGAGGTTTCT 1724
DB 2736 CAGAGCTGAGAGAGTCAAGCTTCAATGAGTTTCAAGAGAAAGAGATGAGAGGTTTCT 2795
QY 1735 GGAAGAGAGAGAGTCTGATTAAGATCAAGAGAAAGAGATGAGAGATGAGAGAG 1784
DB 2796 GGAAGAGAGAGAGTCTGATTAAGATCAAGAGAAAGAGATGAGAGATGAGAGAG 2855
QY 1785 GATATCAGAGAGATTTTGAATTTGATGAGAGAAAGATTTGATGAGGCTTTGAGACGCTCAT 1844
DB 2856 GATATCAGAGAGATTTTGAATTTGATGAGAGAAAGATTTGATGAGGCTTTGAGACGCTCAT 2915

```

```

QY 1845 GTACAAGCATGGCCTTCAACATGAAGATGATTGA 1878
DB 2916 GTACAAGCATGGCCTTCAACATGAAGATGATTGA 2949

RESULT 3
ABX81661
ID ABX81661 standard; cDNA; 272 BP.
XX
AC ABX81661;
XX
DT 24-APR-2003 (first entry)
XX
DE Corn ear-derived polynucleotide (cpd) #121.
XX
KW Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW structural gene; functional gene; regulatory gene;
KW corn ear-specific profile; gene transcription; gene expression;
KW hybrid plant; desirable trait expression; plant breeding program;
KW inheritance; desirable characteristic; growth; development;
KW disease resistance; environmental adaptability; quality; yield;
KW multigene trait; plant; gene; ss.
XX
OS Zea mays.
XX
PN US6476212-B1.
XX
PD 05-NOV-2002.
XX
PF 14-MAY-1999; 99US-00313294.
XX
PR 26-MAY-1998; 98US-0086722P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lalundi RV, Ito LY, Sherman BK;
XX
DR MPI; 2003-208840/20.
XX
PT Novel purified corn-ear derived polynucleotide useful as hybridization
PT probe for detecting polynucleotide in sample, and for identifying,
PT evaluating, and altering desired characteristics associated with growth,
PT development.
XX
PS Example; SEQ ID NO 121; 390bp; English.
XX
CC The present invention relates to the isolation of corn ear-derived
CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC and SATMON023. Some of the cdps uniquely identify structural, functional,
CC and regulatory genes of corn ear. The polynucleotides sequences are
CC useful for detecting cdps in a sample, for producing a corn ear-specific
CC profile of gene transcription, for detecting altered gene expression in
CC inbred or hybrid plants, and for screening several molecules for specific
CC binding to the polynucleotide. The cdps are useful to identify, isolate,
CC or extend identical or related corn-ear nucleic acid sequences from DNA
CC libraries, and in nucleic acid amplification or hybridisation techniques
CC to follow the expression of desirable traits through plant breeding
CC program. Preferably, the cdps are used to identify, evaluate, alter, or
CC follow the inheritance of desired characteristics associated with growth
CC and development, disease resistance, environmental adaptability, quality,
CC and yield of corn. The cdps are also useful as molecular markers for
CC studying inheritance and multigene traits in a plant breeding program.
CC The cdps are useful for producing purified corn-ear polypeptides by
CC recombinant techniques. They are also useful in diagnostic assays to
CC detect or confirm conditions or diseases associated with abnormal levels
CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC polynucleotides (cdps) of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/psipsdIDEntry.html
XX
SQ Sequence 272 BP; 79 A; 52 C; 68 G; 67 T; 0 U; 6 Other;

```

Query Match 4.0%; Score 75; DB 7; Length 272;
 Best Local Similarity 57.6%; Pred. No. 8.7e-11;
 Matches 148; Conservative 0; Mismatches 106; Indels 3; Gaps 1;

QY 1009 TATGAGGCTTTAGAGCAGCCATTCTTATGTCCACAGGGCCATGTGGAGTGT 1068
 DB 2 TATGAAGCAAGTAAACACACCTCATGCTTATGTCCTCTGGGACCGGTGTATGACG 61
 QY 1069 CTGATGTTTGAAGCAGTGCCTGCTATTTGAGGCGGAGCCCTCCACGGGAGTTA 1128
 DB 62 TTAATATTTGAAGCTCAGCTGTGGGTACATGAAAGCTGAAGCTGTGCTAAACACTTT 121
 QY 1129 GCTGAGATGGGTTAGATGAATTTGCTG3---GCTCAGAAGCGCAGTATGTTTCTGA 1185
 DB 122 GTTAAACAGATGACACAGAGATTCATGSCACCTACGCGANGTTGATTTGGCTGCT 181
 QY 1186 GGTGTTGCGCAACTGATGCTTCTTGCACGAGCAAGATGTGACATATTCATCA 1245
 DB 182 GGGAAAGCGCCTATATGTTTCTTAGCAACAAAGAGATATGACGATTTAAACAAG 241

QY 1246 CACTCTCAAGGCAAAAC 1262
 DB 242 CAGCCATGATTAAGC 258

RESULT 4
 ABX82003
 ID ABX82003 standard; cDNA; 274 BP.
 XX ABX82003;
 DT 24-APR-2003 (first entry)
 XX
 DE Corn ear-derived polynucleotide (cpd) #463.
 XX
 XX Corn ear-derived polynucleotide; cpd; cDNA library; SATMON022; SATMON023;
 KM structural gene; functional gene; regulatory gene;
 KM corn ear-specific profile; gene transcription; gene expression;
 KM hybrid plant; desirable trait expression; plant breeding program;
 KM inheritance; desired characteristics; growth; development;
 KM disease resistance; environmental adaptability; quality; yield;
 KM multigene trait; plant; gene; ss.
 OS
 XX
 XX Zea mays.
 OS
 XX
 XX US6476212-B1.
 XX
 XX 05-NOV-2002.
 XX
 XX 14-MAY-1999; 99US-00313294.
 XX
 XX 26-MAY-1998; 98US-0086722P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Lalgudi RV, Ito LY, Sherman BK;
 DR WPI, 2003-20840/20.
 XX
 XX Novel purified corn-ear derived polynucleotide useful as hybridization
 PT probe for detecting polynucleotide in sample, and for identifying,
 PT evaluating, and altering desired characteristics associated with growth,
 PT development.
 XX
 XX Example; SEQ ID NO 463; 390PP; English.
 XX
 CC The present invention relates to the isolation of corn ear-derived
 CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
 CC and SATMON023. Some of the cdps uniquely identify structural, functional,
 CC and regulatory genes of corn ear. The polynucleotides sequences are
 CC useful for detecting cdps in a sample, for producing a corn ear-specific
 CC profile of gene transcription, for detecting altered gene expression in
 CC inbred or hybrid plants, and for screening several molecules for specific

CC binding to the polynucleotide. The cdps are useful to identify, isolate,
 CC or extend identical or related corn-ear nucleic acid sequences from DNA
 CC libraries, and in nucleic acid amplification or hybridization techniques
 CC to follow the expression of desirable traits through plant breeding
 CC programs. Preferably, the cdps are used to identify, evaluate, alter, or
 CC follow the inheritance of desired characteristics associated with growth
 CC and development, disease resistance, environmental adaptability, quality,
 CC and yield of corn. The cdps are also useful as molecular markers for
 CC studying inheritance and multigene traits in a plant breeding program.
 CC The cdps are useful for producing purified corn-ear polypeptides by
 CC recombinant techniques. They are also useful in diagnostic assays to
 CC detect or confirm conditions or diseases associated with abnormal levels
 CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
 CC polynucleotides (cdps) of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/psipdsIdentify.html
 XX
 XX
 SQ Sequence 274 BP; 83 A; 49 C; 73 G; 67 T; 0 U; 2 Other;

Query Match 4.0%; Score 74.4; DB 7; Length 274;
 Best Local Similarity 60.5%; Pred. No. 1.3e-10;
 Matches 156; Conservative 0; Mismatches 98; Indels 4; Gaps 2;

QY 1009 TATGAGGCTTTAGAGCAGCCATTCTTATGCT-CCACAGGGCCATGTGGATGAGTGT 1067
 DB 2 TATGAAGCAAGTAAACACACCTCATGCTTATGTCCTCTGGGACCGGTGTATGACG 61
 QY 1068 TCTGATGTTTGAAGCAGTGCCTGCTATTTGAGGCCGGAACGCTTCCACCGGAGTT 1127
 DB 62 GTTAAATATTTGAAGCTCAGCTGTGGGTACATGTAAGCTGAACCTCTCATTAACACTT 121
 QY 1128 AGCTGATGGGGGTTAGATGAATTTGCTG3---GGTCAAGAGCGCAGTATGTTTCTG 1184
 DB 122 TGTATATCAAGGTACAGACAGAGATTCATGACCTTACGCAAGTTTGATTTGCTG 181
 QY 1185 AGGTGTTGCGCAACTGATGCTTCTTGCACGAGCAAGATGTGACATATTCATCA 1244
 DB 182 TGGGAAAGCAAGCTATATGTTTCTTAGCAACAAAGAGATATGAGGCAATTTAAACA 241

QY 1245 ACACTCTCAAGGCAAAAC 1262
 DB 242 GCATTGCCAGGTAAGC 259

RESULT 5
 ABL71922
 ID ABL71922 standard; cDNA; 310 BP.
 XX ABL71922;
 DT 14-MAY-2002 (first entry)
 XX
 XX Corn tassels-derived polynucleotide (cdps) SEQ ID NO:1296.
 DE
 XX
 XX Corn; corn tassels-derived polynucleotide; cdps; hybrid breeding; CDPS;
 KM inheritance; characteristic; growth; development; disease resistance;
 KM environmental adaptability; quality; yield; molecular marker;
 KM multigene trait; plant breeding; corn tassels; gene; ss.
 OS
 XX
 XX Zea mays.
 OS
 XX
 XX US2001051335-A1.
 XX
 XX 13-DEC-2001.
 XX
 XX 16-APR-1999; 99US-00294093.
 XX
 XX 21-APR-1998; 98US-0082567P.
 XX
 XX (IAGC/) LALGUDI R V,
 PA (ITOL/) ITO L Y,
 PA (SHER/) SHERMAN B K.

QY 1278 GTTGAATCATACCAAGATGGTTGTAAGAGCTGAGCAGATCTCTGAGACAATCA 1337
 DB 1860 GCAGCAGATGACACAGACAGATGACGACGACGATGACAGACAGATGACGA 1919
 QY 1338 GCAGCTGAATCTTTAAGAACAGCTCTCAAAACAGAACAGACAGCCAGGTCCTGA 1397
 DB 1920 GCAGCAGATGACACAGACAGATGACGACGACGATGACAGACAGATGACGA 1979
 QY 1398 GGAATCTCTGAAATTTATGACGACAAAGCTGCTGAACCTGACAGATTAATCGATCGT 1457
 DB 1980 GCAGCAGATGACACAGACAGATGACGACGACGATGACAGACAGATGACGA 2039
 QY 1458 GAGACAGAGAACTTAAGATGACATGAAACAGAACAGGAGATGATCAGACAGCAG 1517
 DB 2040 GCAGCAGATGACACAGACAGATGACGACGACGATGACAGACAGATGACGA 2099
 QY 1518 GTTTTCATGATTCATCAATCAACAGATCCATGAAAGAACAGCAAGAGAGAAATTT 1577
 DB 2100 GCAGCAGATGACACAGACAGATGACGACGACGATGACAGACAGATGACGA 2159
 QY 1578 CGAGATGTCGACACAGACAGATGTCGCAAGGTTGTCGACAGACAGCAAACTTAA 1637
 DB 2160 GGATGACGACGACAGACAGATGACGACGACGATGACAGACAGATGACGA 2219
 QY 1638 TCCCTCTAGCAATGACGATTCGCCAAAGAGCTGAGAAAGTGTCAAGCTTCAGGATT 1697
 DB 2220 GCAGCAGACGACGATGACGACGACGACGATGACGACGACGATGACGACGACGA 2279
 QY 1698 TCAAGAGAAAGATGAGAGATGTTGTGAAGAGAGAGATGCTGATAAAGATCAAGA 1757
 DB 2280 GCAGCAGAGAGACAGACAGAGATGTAAGAGACGACGACAGAGATGAGATCAGGA 2339
 QY 1758 GAAGAAGATGAAAGACATGAAGAAGAGCATCCAGAGAGATTTTGTGAGAAAGA 1817
 DB 2340 GCAGGAGTTAGAGACAGACAGAGATTTAGAGAGACGACGACAGAGATTTAGAGAGCA 2399
 QY 1818 ATTGATGAGGCTTTGGAACAG 1839
 DB 2400 GCAGCAGAGTTAGAGAGCAG 2421
 RESULT 7
 AAF82901
 ID AAF82901 standard; DNA; 3489 BP.
 AC AAF82901;
 XX
 DT 11-SEP-2003 (revised)
 DT 06-AUG-2003 (revised)
 DT 29-JUN-2001 (first entry)
 XX
 DE Nucleotide sequence of KSHV tethering protein, LANA.
 XX
 KM Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
 KM Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
 KM KSHV; latency-associated nuclear antigen; LANA; ds.
 XX
 OS Human herpesvirus 8.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3489
 FT /*tag= a
 XX
 PN WO200125484-A2.
 XX
 PD 12-APR-2001.
 XX
 PF 29-SEP-2000; 2000MO-US026908.
 XX
 PR 01-OCT-1999; 99US-00410399.
 XX
 PA (UNMI) UNIV MICHIGAN.

XX
 PI Robertson ES, Colter MA;
 XX
 DR WPI; 2001-281736/29.
 DR P-PSDB; AAF82331.
 XX
 PT A composition for use in gene therapy comprises an expression vector that
 XX includes a nucleic acid sequence encoding a nucleic acid binding protein.
 PS Disclosure; Fig 9A; 60pp; English.
 XX
 CC The invention provides a composition comprising nucleic acid, histone H1
 CC protein and expression vector operationally encoding a protein suitable
 CC for tethering the nucleic acid to the histone H1 protein, where the
 CC tethering protein is LANA. The composition is useful in aiding the
 CC retention of the viral DNA in the host cell. The viral vector encodes a
 CC protein suitable for tethering DNA to histone H1. Methods for screening
 CC for compounds which are agonistic or antagonistic for the tethering of
 CC viral proteins to histone H1 and DNA binding sites are useful for
 CC developing the method of viral transfer. The composition has applications
 CC to gene therapy, including the treatment of multiple sclerosis,
 CC Parkinson's disease, Huntington disease and diabetes. The present
 CC sequence represents the nucleotide sequence of the Kaposi's sarcoma
 CC associated herpesvirus (human herpesvirus 8) latency-associated nuclear
 CC antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-
 CC 2003 to correct OS field.) (Updated on 11-SEP-2003 to standardise OS
 CC field)
 XX
 SQ Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;
 Query Match 3.1%; Score 58.8; DB 4; Length 3489;
 Best Local Similarity 43.4%; Pred. No. 1.4e-05;
 Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;
 QY 1218 GAAGCAAGATCTGACATATTCATCAATCACTCTCAAGCAAAACAGGCTGAATTCGA 1277
 DB 1800 GCAGCAGATGACGACGACAGATGACGACGACGATGACGACGACGATGACGA 1859
 QY 1278 GTTGAATCATACCAAGATGTTGTGAAGAGATGAGGATCTGAGACAAATCA 1337
 DB 1860 GCAGCAGATGACGACGACAGATGACGACGACGATGACGACGACGATGACGA 1919
 QY 1338 GCAGCTGAATCTTTAAGAACAGCTCTCAAAACAGAACAGCCAGGTCCTGA 1397
 DB 1920 GCAGCAGATGACACAGACAGATGACGACGACGATGACAGACAGATGACGA 1979
 QY 1398 GGAATCTCTGAAATTTATGACGACAAAGCTGCTGAACCTGACAGATTAATCGATCGT 1457
 DB 1980 GCAGCAGATGACACAGACAGATGACGACGACGATGACAGACAGATGACGA 2039
 QY 1458 GAGACAGAGAACTTAAGATGACATGAAACAGAACAGGAGATGATCAGACAGCAG 1517
 DB 2040 GCAGCAGATGACACAGACAGATGACGACGACGATGACAGACAGATGACGA 2099
 QY 1518 GTTTTCATGATTCATCAATCAACAGATCCATGAAAGAACAGCAAGAGAGAAATTT 1577
 DB 2100 GCAGCAGATGACACAGACAGATGACGACGACGATGACAGACAGATGACGA 2159
 QY 1578 CGAGATGTCGACACAGACAGATGTCGCAAGGTTGTCGACAGACAGCAAACTTAA 1637
 DB 2160 GGATGACGACGACAGACAGATGACGACGACGATGACAGACAGATGACGA 2219
 QY 1638 TCCCTCTAGCAATGACGATTCGCCAAAGAGCTGAGAAAGTGTCAAGCTTCAGGATT 1697
 DB 2220 GCAGCAGACGACGATGACGACGACGACGATGACGACGACGATGACGACGACGA 2279
 QY 1698 TCAAGAGAAAGATGAGAGATGTTGTGAAGAGAGAGATGCTGATAAAGATCAAGA 1757
 DB 2280 GCAGCAGAGAGACAGACAGAGATGTAAGAGACGACGACAGAGATGAGATCAGGA 2339
 QY 1758 GAAGAAGATGAAAGACATGAAGAAGAGCATCCAGAGAGATTTTGTGAGAAAGA 1817
 DB 2340 GCAGGAGTTAGAGACAGACAGAGATTTAGAGAGACGACGACAGAGATTTAGAGAGCA 2399

QY 1818 ATTGATGAGCGCTTGGAAACAG 1839
 DB 2400 GGAGCAGAGTGAAGAGCAGCAG 2421

RESULT 8
 ABA93487
 ID ABA93487 standard; DNA; 3489 BP.
 XX ABA93487;
 AC
 XX
 XX 25-APR-2002 (first entry)
 DT
 XX
 XX Kaposi's sarcoma-associated herpesvirus LANA protein encoding DNA.
 DE
 XX Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RYCAB; PEL;
 KM KSHV terminal repeat; rhadino virus cis acting element; episome;
 KW primary effusion lymphoma; latency-associated nuclear antigen;
 XX gene therapy; gene transfer; gene; ds.
 XX
 OS Human herpesvirus 8.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3489
 FT /*tag= a
 FT /product= "LANA protein"
 FT /note= "latency-associated nuclear antigen"

XX US6322792-B1.
 XX 27-NOV-2001.
 PD
 XX 21-APR-1999; 99US-00298568.
 XX
 XX 19-NOV-1998; 98US-0109422P.
 PR
 XX (KIEF/) KIEFF E D.
 PA
 XX Kieff ED, Balleskas ME, Kaye KM;
 PI WPI: 2002-153769/20.
 XX P-PSDB; ABB05621.
 DR
 XX
 XX System for episomal retention of plasmids in mammalian cells, useful in
 PT gene therapy, comprises rhadinoviral LANA and RVCAB sequences.
 PS
 XX Claim 1; Fig 6; 27pp; English.

XX The present invention describes a system (A) for maintaining a plasmid as
 CC an episome in mammalian cells, comprising the rhadinoviral sequence LANA
 CC (latency-associated nuclear antigen) of 3489 base pairs (see ABA93487,
 CC S1) expressed in the cell, and the rhadinoviral sequence RVCAB
 CC (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, S2)
 CC present in the plasmid. Also describes a method for maintaining a
 CC closed circular DNA in a cell by expressing (S1) in the cells and having
 CC (S2) as a cis-acting and maintenance sequence in the DNA. (A) is
 CC particularly used in gene therapy (or other gene transfer applications)
 CC that uses mammalian cells in which LANA is expressed. (A) improves
 CC persistence of gene therapy vectors in cells. The present sequence
 CC encodes Kaposi's sarcoma-associated herpesvirus (KSHV, also called human
 CC herpesvirus 8) LANA protein, which is used in the exemplification of the
 CC present invention

QY Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;
 DB
 Query Match 3.1%; Score 58.8; DB 6; Length 3489;
 Best Local Similarity 43.4%; Pred. No. 1.4e-05;
 Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

QY 1218 GAAGCAAGATCTGACATATTCATCAACACTCTTCAAGGCAAAACAGGCTGAATTCGA 1277
 DB 1800 GCAGCAGATGAGCAGCAGCAGATGACAGCAGCAGCAGATGACAGCAGCAGATGAGCA 1859

QY 1278 GTTGAAATCATACCAAGAGATGTTTAAAGAGCTGAGCAGATCTTGAGCAATCA 1317
 DB 1860 GCAGCAGATGAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGATGAGCA 1919

QY 1338 GCAGCTGAATCTTAAAGAAACAGCTCTCAAAAAGAAACAGCAGCCAAAGTCTTGA 1397
 DB 1920 GCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGATGAGCA 1979

QY 1398 GGAATCTTGAAATTTATGAGCCGAAAGCTGCTGTAAGTGCAGATTAATCGATCGT 1457
 DB 1980 GCAGCAGATGAGCAGCAGCAGATGAGCAGCAGCAGCAGCAGATGAGCAGCAGATGAGCA 2039

QY 1458 GAGACAGAACTAAGATGACATGAAACAGAAACAGGAAAGATGATGACACAGCAAG 1517
 DB 2040 GCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGATGAGCA 2099

QY 1518 GTTTTTCATGATTCATCAATCAACAGATCCATGAAAGAAAGAGCCAAAGAGATTT 1577
 DB 2100 GGAGCAGCAGATGAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGATGAGCA 2159

QY 1578 CGAGATGTCAGCAGCAGCAGCAAGCTGCCAAGCTTTGCGCAGCAGCAGCAAACTTAA 1637
 DB 2160 GGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGCAGATGAGCAGCAGCAGATGAGCA 2219

QY 1638 TCCCTCTAGCAATGACGATGCGCAAGAGAGCTGAGGAAGTGTCAAGCTTCATCGAGTT 1697
 DB 2220 GCAGCAGCAGCAGATGAGCAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCA 2279

QY 1698 TCAGAGAAAGAGATGAGAGAGATTGTGAAAGAGAGGAGATCTGATTAAGATCAAGA 1757
 DB 2280 GCAGCAGGAGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGAGATGAGCAGCAGCA 2339

QY 1758 GAAGAAATGAGAAACATGAAGAAGAGCATACAGAGAGATTTGATCTGAGAAAGA 1817
 DB 2340 GCAGAGTTAGAGAGCAGCAGCAGAGATGAGAGAGCAGCAGAGATGAGAGCA 2399

QY 1818 ATTGATGAGCGCTTGGAAACAG 1839
 DB 2400 GGAGCAGAGTGAAGAGCAGCAG 2421

RESULT 9
 AAV73805/c
 ID AAV73805 standard; DNA; 32207 BP.
 XX AAV73805;
 AC
 XX 17-OCT-2003 (revised)
 DT 25-FEB-1999 (first entry)
 XX
 DE KSHV LTR DNA (nucleotides 105,301-137,507).

XX Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;
 KM dihydrofolate reductase; LTR; long unique region; vaccine; prophylaxis;
 KM diagnosis; treatment; HHV8; capsid protein IV; tegument protein IV;
 KM glycoprotein; kaposin; cyclin D; immediate early protein; IE; OX-2;
 KW v-adh; G-protein coupled receptor; FGFR4; ds.

XX Human herpesvirus 8.
 OS
 XX US5849564-A.
 XX 15-DEC-1998.
 PD
 XX 29-NOV-1996; 96US-00770379.
 PF
 XX 29-NOV-1996; 96US-00770379.
 PR
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 PA
 XX Bohenzky RA, Moore PS, Ruseo JU, Chang Y, Edelman IS;
 PI

DR WPI; 1999-069741/06.
 XX Kaposi's sarcoma-associated herpes virus nucleic acid - encodes
 PT di:hydrofolate reductase and is useful for treatment, prophylaxis or
 PT diagnosis of Kaposi's sarcoma.
 XX
 XX Disclousure; Col 155-182; 109pp; English.
 PS
 CC This sequence is a fragment of the Kaposi's sarcoma-associated
 CC herpesvirus (KSHV) LIR (long unique region). This fragment contains
 CC coding regions for ORF65 which encodes capsid protein IV, ORF66, ORF67
 CC which encodes tegument protein IV, ORF68 which encodes a glycoprotein,
 CC ORF69, K12 which encodes Kaposin, K13, ORF72 which encodes cyclin D,
 CC ORF73 which encodes immediate early protein (IEP), K14 which encodes OX-2
 CC (v-adr), ORF74 which encodes G-protein coupled receptor, ORF75 which
 CC encodes tegument protein/FGAART, K15. KSHV is a new human Herpesvirus
 CC (HHV8) believed to cause Kaposi's sarcoma (KS) which is the most common
 CC form of neoplasm occurring in persons with acquired immune deficiency
 CC syndrome (AIDS). The DHFR protein is useful for vaccination, prophylaxis,
 CC diagnosis and treatment of a subject with Kaposi's sarcoma and for
 CC detecting expression of a DNA virus associated with Kaposi's sarcoma in a
 CC cell. (updated on 17-OCT-2003 to standardise OS field)
 CC
 XX
 XX Sequence 32207 BP; 7229 A; 9156 C; 8713 G; 7109 T; 0 U; 0 Other;
 SQ
 Query Match 3.1%; Score 58.8; DB 2; Length 32207;
 Best Local Similarity 43.4%; Pred. No. 4.5e-05;
 Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;
 QY 1218 GAAGCAAGATCTGGACATATTCAATCAACCTCTCAAGGCAAAACAAGCTGAATTCGA 1277
 Db 20197 GCAGCAGATGACAGCAGCAGATGATGACGACGAGATGACAGCAGATGACG 20138
 QY 1278 GTTGAATATCAACAGAGATGTTGTAAGAGCTGAGGCAATCTCTGAGGACATCA 1337
 Db 20137 GCAGCAGATGACAGCAGCAGATGATGACGACGAGATGACAGCAGATGACG 20078
 QY 1338 GCAGTGAATCTTAAAGAACAGCTCTCAAAACGAACACAGCAGCCAGGTCCTTGA 1397
 Db 20077 GCAGCAGATGACAGCAGCAGATGATGACGACGAGATGACAGCAGATGACG 20018
 QY 1398 GGAATCTCTGAAATTTATGACGAGAACTGCTGTAAGATCTGAGGATTAATCGATCGT 1457
 Db 20017 GCAGCAGATGACAGCAGCAGATGATGACGACGAGATGACAGCAGATGACG 19958
 QY 1458 GACACGAGAACTTAAGATGACATGAACAAGAGGAAAGATGATGACACACG 1517
 Db 19957 GCAGCAGATGACAGCAGCAGATGATGACGACGAGATGACAGCAGATGACG 19898
 QY 1518 GTTTTCATGATTCATCAATCAACAGATCCATGTAAGAAAGAGACGCAAGGAGGAATTT 1577
 Db 19897 GCAGCAGCAGATGACAGCAGCAGATGATGACGACGAGATGACAGCAGATGACG 19838
 QY 1578 CGAGATGTTGACAGCAGCAGAGAACTGCAAGTTGTTGGCCAGCAGCAGCAATTA 1637
 Db 19837 GATGACGACAGCAGCAGATGATGACGACGAGATGACAGCAGATGACG 19778
 QY 1638 TCCCTTAGCAATGACGATTTGCCGAAAGAGAGCTGAGAAAGTCTCAAGCTTCA 1697
 Db 19777 GCAGCAGCAGCAGATGACAGCAGCAGATGAAACAGAGACAGCAGAGAGAGCA 19718
 QY 1698 TCAGAGAAAGAGATGAGAGAGTTTGTGAAGAGAGAGAGATCTGATAAAGATCAAG 1757
 Db 19717 GCAGCAGAGAGCAGAGAGAGATTAAGAGAGCAGAGCAGAGATTAAGATCAAG 19658
 QY 1758 GAAAGAGATGAAAGACATGAAGAGGACATCAAGAGAGATTAATTTGATCTGAGAAAG 1817
 Db 19657 GCAGAGATTAGAAGCAGAGAGAGATTAGAAGAGCAGAGAGATTAAGAGAGCA 19598
 QY 1818 ATTGATGAGGCTTTGGAACG 1839
 Db 19597 GGAGCAGAGATTAGAAGAGCAG 19576

RESULT 10
 AAV19941/C
 ID AAV19941 standard; DNA; 137507 BP.
 XX
 XX AAV19941;
 AC
 XX 27-AUG-2003 (revised)
 DT 03-AUG-1998 (first entry)
 XX
 DE KSHV long unique coding region and terminal repeat.
 XX
 XX KSHV; HHV8; human herpes virus 8; macrophage inflammatory protein II;
 XX interleukin-6; IL-6; interferon regulatory factor; rheumatoid arthritis;
 XX complement-binding protein; glycoprotein; capsid protein IV; infection;
 XX immediate early protein; Kaposi's sarcoma; protective vaccine; lymphoma;
 XX lymphoproliferative disease; leukaemia; splenomegaly; mycosis fungoides;
 XX HIV immune status; anti-inflammatory agent; therapy; ds.
 XX
 OS Human herpesvirus 8.
 XX
 FH Location/Qualifiers
 FT 1142..2794
 FT /tag= a
 FT /product= "complement-binding protein"
 FT CDS
 FT /tag= b
 FT /product= "glycoprotein B"
 FT CDS
 FT /tag= c
 FT /product= "interleukin 6"
 FT CDS
 FT /tag= d
 FT /product= "macrophage inflammatory protein II"
 FT CDS
 FT /tag= e
 FT /product= "interferon regulatory factor 1"
 FT CDS
 FT /tag= f
 FT /product= "protein T1.1"
 FT CDS
 FT /tag= g
 FT /product= "glycoprotein M"
 FT CDS
 FT /tag= h
 FT /product= "glycoprotein L"
 FT CDS
 FT /tag= i
 FT /product= "interferon regulatory factor 2"
 FT CDS
 FT /tag= j
 FT /product= "interferon regulatory factor 3"
 FT CDS
 FT /tag= k
 FT /product= "glycoprotein X"
 FT CDS
 FT /tag= l
 FT /product= "interferon regulatory factor 4"
 FT CDS
 FT /tag= m
 FT /product= "capsid protein IV"
 FT CDS
 FT /tag= n
 FT /product= "immediate early protein"
 XX
 PN W09804576-A1.
 XX
 XX 05-FEB-1998.
 PD
 XX 22-JUL-1997;
 PF 97WO-US013346.
 XX
 XX 25-JUL-1996;
 PR 96US-00686243.
 XX
 XX 25-JUL-1996;
 PR 96US-00686349.

PR 25-JUL-1996; 96US-00686350.
 PR 25-JUL-1996; 96US-00687253.
 PR 05-SEP-1996; 96US-00708678.
 PR 10-OCT-1996; 96US-00728323.
 PR 13-NOV-1996; 96US-00747887.
 PR 13-NOV-1996; 96US-00748640.
 PR 29-NOV-1996; 96US-00757669.
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX PA
 XX PI Chang Y, Bohenzky RA, Russe JJ, Edelman IS, Moore PS;
 XX DR WPI, 1998-130615/12.
 XX
 PT New nucleic acid encoding Kaposi's sarcoma associated herpes virus
 PT proteins - useful for, e.g. detecting levels of HHV8 In, and preparation
 PT of vaccines for treatment of, HIV patients.
 XX
 XX Example 2; Page 135-203; 230pp; English.

CC This sequence represents the long unique region and terminal repeat of
 CC the Kaposi's sarcoma-associated herpes virus (KSHV). KSHV is also known
 CC as human herpes virus 8 (HHV8). This sequence contains the DNAs of the
 CC invention which encode KSHV polypeptides selected from: (a) viral
 CC macrophage inflammatory protein (MIP) II; (b) viral interleukin-6 (IL-6);
 CC (c) viral IRF 1; (d) complement-binding protein; glycoproteins B, M or L;
 CC (d) capsid protein IV encoded by ORF65; and (e) immediate early protein
 CC encoded by ORF73. Labelled probes for the nucleic acid, proteins encoded
 CC by it, and antibodies (Ab) specific for the proteins are useful for
 CC detecting HHV8, specifically for diagnosis of Kaposi's sarcoma, in body
 CC fluids or tissue samples. HHV8 infections can be treated with antisease
 CC or triplex forming molecules or agents that bind specifically to the
 CC protein. Ab may be used for prophylaxis or treatment of HHV8 infection,
 CC while the protein can be used in protective vaccines. Ab may also be used
 CC to differentiate between lymphomas, and HHV8 may be implicated in many
 CC other lymphoproliferative diseases such as lymphomas, leukaemia,
 CC splenomegaly and mycosis fungoides. Cells and animals containing the
 CC nucleic acid are useful for drug screening. HHV8-derived peptides can be
 CC used as targets for antiviral drugs, e.g. dihydrofolate reductase gene
 CC can be inhibited with methotrexate. These can also be used to determine
 CC the immune status of a patient infected with HIV. HHV8 derived protein
 CC viral MIP III may be used as an anti-inflammatory agent for, e.g.
 CC treating rheumatoid arthritis. This sequence is stated as containing 81
 CC open reading frames. (Updated on 27-AUG-2003 to correct OS field.)
 XX
 XX Sequence 137507 BP; 32579 A; 37795 C; 35758 G; 31375 T; 0 U; 0 Other;

Query Match 3.1%; Score 58.8; DB 2; Length 137507;
 Best Local Similarity 43.4%; Pred. No. 9.7e-05;
 Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

QY 1218 GAAGCAAGATCTGGACATATTAATCAATCACTCTCAAGGAAACAAAGGCTGAATTCGA 1277
 DB 125497 GCAGCAGATGACGACGACGACGATGACGACGATGACGACGATGACGACGATGACGACG 125438
 QY 1278 GTTGAATATATACCAAGATGTTTAAAGAGCTGAGGAGATCTCTGAGGACATCA 1337
 DB 125437 GCAGCAGATGACGACGACGATGACGACGATGACGACGATGACGACGATGACGACG 125378
 QY 1338 GCAGCTGAATCTTTAAGAACAGCTCTCAAAACGAACAGCAGCCAGGCTGTTGA 1397
 DB 125377 GCAGCAGATGACGACGACGATGACGACGATGACGACGATGACGACGATGACGACG 125318
 QY 1398 GGAATCTCGAATTAATGAGGAGAAAGCTGCTAAGCTGAGAGATTAATGATCGT 1457
 DB 125317 GCAGCAGATGACGACGACGATGACGACGATGACGACGATGACGACGATGACGACG 125258
 QY 1458 GAGACAGAACTTAAGATGACGATGAACAGAAACAGGAAGAGATGATGACACAGACAG 1517
 DB 125257 GCAGCAGATGACGACGACGATGACGACGATGACGACGATGACGACGATGACGACG 125198
 QY 1518 GTTTTCATGATTCATCAACAGATCCATGAAAGAGAGCGAAAGGAGGAAATTT 1577

DB 125197 GGAGCAGCAGATGACGACGACGACGATGACGACGACGATGACGACGATGACGACGACG 125138
 QY 1578 CGAGATGTTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1637
 DB 125137 GATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 125078
 QY 1638 TCCCTCTAGCAATGACGATTTGCCGAAAGAGACCTGAGGAAGTGTCAAGCTTCATGAGTT 1697
 DB 125077 GCAGCAGCAGCAGATGACGACGACGACGACGATGAACAGGACGACGACGACGACG 125018
 QY 1698 TCAGAGAAAGAGATGAGAGATTTGTGGAAGAGGAGATGCTGATTAAGATCAAG 1757
 DB 125017 GCAGCAGGAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 124958
 QY 1758 GAAGAAATGAGAGACATGAAAGAGGACATGACGACGACGATATTTGATCTGGAAGAA 1817
 DB 124957 GCAGGAGTTAGAGGACGACGACGACGACGACGATTAAGAGACGACGACGACGACG 124898
 QY 1818 ATTTGATGAGGCTTTGGAACAG 1839
 DB 124897 GCAGCAGGAGTTAGAGGACGACG 124876
 RESULT 11
 ID AB199537 standard; cDNA; 1080 BP.
 XX AB199537;
 AC AB199537;
 XX 07-MAR-2002 (first entry)
 DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:537.
 XX
 KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
 OS Mus musculus.
 XX WO200181818-A2.
 PN 22-NOV-2001.
 PD 18-MAY-2001; 2001WO-0P004192.
 PF 18-MAY-2001; 2000JP-00145977.
 PR 18-MAY-2000; 2000JP-00145977.
 PA (UNIV-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 PI P-PSDB; ABB57221.
 XX WPI; 2002-034733/04.
 PS Claim 2; Page 1472-1473; 2690pp; English.
 XX
 CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring the
 CC expression levels of particular genes (AB199202 to AB199912, encoding the
 CC protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The expression
 CC levels or expression profiles produced by these genes are used as an
 CC indicator when screening for ischaemic condition-improving drugs or
 CC therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR
 CC primers for a mouse ischaemic condition related sequence, which are used

CC in the exemplification of the present invention
XX Sequence 1080 BP; 370 A; 191 C; 390 G; 129 T; 0 U; 0 Other;

Query Match 3.0%; Score 56.2; DB 6; Length 1080;
Best Local Similarity 48.0%; Pred. No. 4.1e-05;
Matches 160; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 1466 GAACTAAGATGACATGAAACAGACAGGAAAGTGTGATGACACGACAGGTTTTC 1525
DB 701 GAAAG 760
QY 1526 TGGATTCATCAAAACATGATCATGAAAGAGAGAGAGAGAGAGAGATTTTGA 1585
DB 761 AG 820
QY 1586 TGCAGACAG 1645
DB 821 AG 880
QY 1646 GCAATGACGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1705
DB 881 AG 940
QY 1706 AAGAGATGAG 1765
DB 941 AG 1000
QY 1766 TGGAGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1798
DB 1001 AG 1033

RESULT 12

AA087587 standard; DNA; 1686 BP.

AC AA087587;

DT 27-AUG-2003 (revised)
DT 19-DEC-1995 (first entry)

DE DNA encoding Leucocytozoan protozoa structural protein epitope.

KW leucocytozoan protozoa; structural protein; epitope; vaccine; fowl;
leucocytozoanosis; treatment; ss.

OS Leucocytozoan.

PN JP07089995-A.

PD 04-APR-1995.

PF 10-SEP-1993; 93JP-00226078.

PR 10-SEP-1993; 93JP-00226078.

PA (DOHU-) DOHUTSUO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.
(NITS-) NITSEIKEN KK.

DR WPI; 1995-167252/22.
P-PSDB; AAR0491.

PT Immune inducing polypeptide against Leucocytozoan protozoa - useful in
production of vaccines for treatment of Leucocytozoanosis in fowl.

PS Claim 1; Page 12-14; 20pp; Japanese.

CC AA087587-89 encode polypeptides having a whole or partial epitope of a
CC structural protein of Leucocytozoan protozoa (see AAR0491-93). The
CC polypeptides and DNA encoding them are useful in the production of
CC vaccines for the treatment of Leucocytozoanosis of fowl. (Updated on 27-
CC AUG-2003 to correct OS field.)

XX Sequence 1686 BP; 915 A; 129 C; 382 G; 260 T; 0 U; 0 Other;

Query Match 2.9%; Score 55; DB 2; Length 1686;
Best Local Similarity 44.4%; Pred. No. 0.00011;
Matches 220; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 1305 AAAGAGCTGAGGACATCTCTGAGAGACATGACAGCTGAACTTAATTAAGAACAGCT 1364
DB 88 AATGTGAAAG 147
QY 1365 CTCGAAACAGAAACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1424
DB 148 CAAG 207
QY 1425 GCTGCTAGAACTGACAGAGATTAATGATGATGATGATGATGATGATGATG 1484
DB 208 CAAG 267
QY 1485 ACAG 1544
DB 268 GAG 327
QY 1545 CCATGAAAG 1604
DB 328 GAG 387
QY 1605 CAAGGTTGTTGGCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1664
DB 388 AATGAG 447
QY 1665 GAGAGCTGAG 1724
DB 448 GAG 507
QY 1725 GAG 1784
DB 508 GATGAG 567
QY 1785 GCATCAGAGAGAGAT 1799
DB 568 GCTGAGAGAGAGAGAT 582

RESULT 13

AAT91902 standard; DNA; 4000 BP.

AC AAT91902;

DT 07-MAY-1998 (first entry)

DE Mannose-1-phosphate transferase gene MNNA.

KW Mannose-1-phosphate transfer; MNNA gene; enzyme; yeast; regulatory gene;
human; high mannose type neutral saccharide chain; ss.

OS Saccharomyces cerevisiae.

FT Key Location/Qualifiers
FT CDS 418..3953
FT /tag= a
FT /product= "MNNA"

PN JP09266792-A.

PD 14-OCT-1997.

PF 29-MAR-1996; 96JP-00075667.

PR 29-MAR-1996; 96JP-00075667.

PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.

XX WPI; 1997-553460/51.
 DR P-PSDB; AAW30763.
 XX
 PT Positive regulatory gene of mannose-1-phosphate transfer in yeast -
 XX useful for high mannose type neutral saccharide chain production.
 PS Claim 2; Page 14-17; 23pp; Japanese.
 CC This sequence represents the gene of the invention, designated MN4,
 CC encoding a protein which positively regulates mannose-1-phosphate
 CC transfer in yeast. The gene is useful for the preparation of human high
 CC mannose type neutral saccharide chain
 XX
 SQ Sequence 4000 BP; 1453 A; 750 C; 770 G; 1027 T; 0 U; 0 Other;

Query Match 2.9%; Score 55; DB 2; Length 4000;
 Best Local Similarity 47.9%; Pred. No. 0.00018;
 Matches 191; Conservative 0; Mismatches 205; Indels 3; Gaps 1;

QY 1441 GAGATATTCGATCGTGAACAGAGAACTAGATGAGTGAACAGAGAGAG 1500
 Db 3539 GAG 3598
 QY 1501 ATGATGACACAGACAGAGTTTTCATGATTCATCAACAGATCCATGAAAGAGAC 1560
 Db 3599 AAG 3655
 QY 1561 GCAAG 1620
 Db 3656 GAG 3715
 QY 1621 CAGCAG 1680
 Db 3716 AAGAT 3775
 QY 1681 TCAAGCTTCATCGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1740
 Db 3776 GAG 3835
 QY 1741 CTGATTAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1800
 Db 3836 AAGCAG 3895
 QY 1801 TTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1839
 Db 3896 GAAAG 3934

RESULT 14
 AAC40071
 ID AAC40071 standard; DNA; 1140 BP.
 XX
 AC AAC40071;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26940.
 XX
 KM Hybridisation assay; genetic mapping; gene expression control;
 KM protein identification; signal transduction pathway; metabolic pathway;
 KM promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 28-APR-1999; 99US-0130891P.
 PR 30-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 04-MAY-1999; 99US-0132407P.
 PR 05-MAY-1999; 99US-0132484P.
 PR 06-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 07-MAY-1999; 99US-0132487P.
 PR 11-MAY-1999; 99US-0132863P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.

PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158222P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match	Similarity	2.8#;	Score 53.2;	DB 3;	Length 1140;
Best Local	Similarity	45.2#;	Pred. No. 0.0003;		
Matches 196;	Conservative	0;	Mismatches 238;	Indels	Gaps 0;
Qy	1381	CACCCCAAGTGCTCTTGAAGATCTCTGAAATTATGAGCGAGAGTGCCTGAGAACTGCA	1440		
Db	473	CACGAGGATGATGGTGAAGATGCTCACAAAAACATAGTGGTAAAGATTATGATGATGGA	532		
Qy	1441	GAGGATTAATCGGATCGTGAGACAGAGAACTPAAGTSCAGCATGAACAAACAGGAAAGAG	1500		
Db	533	GATGAAGAAGTAAAGAAAGAGAGAAAGAGAAAGAGTAAAGAAAGATGTTAAT	592		
Qy	1501	ATGGAATGACACGACAGGTTTTTATGATTCATCAAACAGATCCATGAAAGAAAGAC	1560		
Db	593	AACCTGAAGATGATGAGTTTAAAGAAAGAAATAGCAGTCAAGAGACATCAT	652		
Qy	1561	GCAAGGAGGGAATTTTCGAGATTTGCAGACAGAGAAAGTCCAGAGTTGTTGGCAG	1620		
Db	653	GATGATGATGTTTATGATGAGAAAGAAAGAAAGAAAGCTTATATGATGATGATGAG	712		
Qy	1621	CAGCAGCAAAACATTATTCCTCTAGCAATGACGATTGCCGAAAGAGAGCTGAGGAAGTG	1680		
Db	713	AAGGAAGAAAGAAAGCACTTATATGATGATGATGATGATGAGAAAGAAAGACATTACAT	772		
Qy	1681	TCAAGCTTCATCGAGTTTCAAGAAAGAGATGAGAGATGTTTGTGAAGAGAGGAGATG	1740		
Db	773	GATGATGATGATGAGAAAGAAAGAAAGAAAGAGATGATGATGATGAGATTAAGAAAGAG	832		
Qy	1741	CTGATTAAGAAGATCAAGAGAAAGATGAGAAACATGAAAGAGGACATCACAGAGAGATA	1800		
Db	833	AAGGAAGCACTTATGATTAATGATGATGATGAGAAAGAAAGAGATTCATCGTGATGAT	892		
Qy	1801	TTTGATCTGGAGAA	1814		
Db	893	GATGATGAGAGAA	906		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 19:25:20 ; Search time 3286.18 Seconds
(without alignments)
17065.771 Million cell updates/sec

Title: US-10-030-829-2

Sequence: 1 atgagcttcaggsgctgctc.....ttccaatgaagatgatga 1878

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pln:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_pro:*
24: em_gse_rod:*
25: em_gse_rod:*
26: em_gse_pbg:*
27: em_gse_vrt:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
C 1	614.4	32.7	693	28 AQ959658 LEREM21TF
C 2	534	28.4	606	28 AQ959657 LEREM21TF
C 3	533	28.4	647	9 AU236368 AU236368
C 4	523.4	27.9	613	28 AQ964580 LERGX20TF

Result No.	Score	Query Length	ID	Description
C 5	513	27.3	619	9 AV566465 AV566465
C 6	511.8	27.3	523	28 AQ959659 LEREM21TF
C 7	468.8	25.0	644	14 CB835509 BN45.045L
C 8	431.2	23.0	708	14 CB815300 BN15.0250
C 9	390	20.8	849	14 CF517786 CAP0005_I
C 10	378.4	20.1	380	9 AV525508 AV525508
C 11	364.4	19.4	808	14 CF517260 CAP0005_I
C 12	359.8	19.2	548	9 A1999551 A1999551
C 13	358.2	19.1	785	14 CA816697 CA816697
C 14	354.8	18.9	697	28 BHS35855 BHS35855
C 15	331.8	17.7	754	14 CF201113 CF201113
C 16	315.8	16.8	753	10 BF003460 BF003460
C 17	297.2	15.8	1159	14 CK210826 CK210826
C 18	296	15.4	457	29 CNS0009H CNS0009H
C 19	289.6	15.4	682	14 CA930182 CA930182
C 20	280.4	14.9	443	28 BH618393 BH618393
C 21	277	14.7	679	13 BO862792 BO862792
C 22	273.2	14.5	704	14 CD939041 CD939041
C 23	268.2	14.3	638	13 BO855967 BO855967
C 24	267	14.2	623	14 CA930229 CA930229
C 25	266.4	14.2	608	14 CA930278 CA930278
C 26	266.2	14.2	675	10 BF646395 BF646395
C 27	265	14.1	709	14 CB879752 CB879752
C 28	255	13.6	724	14 CA262127 CA262127
C 29	254.8	13.6	589	14 CF232247 CF232247
C 30	251.4	13.4	695	12 BG889897 BG889897
C 31	248.4	13.2	855	14 CA264558 CA264558
C 32	245.6	13.1	542	14 CF201172 CF201172
C 33	244	13.0	421	9 AU227277 AU227277
C 34	243.4	13.0	708	14 CA998756 CA998756
C 35	239.2	12.7	747	14 CA239990 CA239990
C 36	237.4	12.6	557	12 BI309649 BI309649
C 37	230.6	12.3	556	13 BO595426 BO595426
C 38	229.8	12.2	635	14 CB937102 CB937102
C 39	229	12.2	290	28 AQ010650 AQ010650
C 40	228.2	12.2	551	14 CF509841 CF509841
C 41	227.2	12.1	565	13 BO779064 BO779064
C 42	223.4	11.9	896	29 CG969741 CG969741
C 43	223.2	11.8	625	14 CA226585 CA226585
C 44	220.6	11.7	506	13 CA004056 CA004056
C 45	217.2	11.6	499	10 AW831280 AW831280

ALIGNMENTS

RESULT 1
LOCUS AQ959658/c
DEFINITION LEREM21TF LRR Arabidopsis thaliana genomic clone LEREM21, genomic survey sequence.
ACCESSION AQ959658
VERSION AQ959658.1 GI:6787359
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsi
REFERENCE
AUTHORS Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Uterbach, T., Felblayum, T., Liang, P., Creasy, T. and Fraser, C.M.
TITLE Genomic survey sequencing of landberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
JOURNAL Unpublished (2000)
COMMENT The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: atc@igr.org
For additional information, see <http://www.tigr.org/tdb/ac/at.html>

Seq primer: TF
Class: shotgun.
Location/Qualifiers

ORIGIN

Query Match	32.7%;	Score	614.4;	DB	28;	Length	693;
Best Local Similarity	99.8%;	Pred.	No. 1.7e-137;				
Matches	615;	Conservative	0;	Mismatches	1;	Indels	0;
						Gaps	0;

QY	134	AGGTATTTCGAGAAAGAACAGAACCAACAGGAAACATCTTGGAATAAACTTGGGTTT	193
Db	688	AGGTATTTCGAGAAAGAACAGAACCAACAGGAAACATCTTGGAATAAACTTGGGTTT	629
QY	194	CTCAGAAATTCGAATCCTCTAGAGCTTGGGTGTGTCAGACGCAAGGAGAGTACCAACG	253
Db	628	CTCAGAAATTCGAATCCTCTAAAGCTTGGGGTGTGCAGACGCAAGGAGAGGTACCAACG	569
QY	254	TATCTGGGAGAGGAAACAAATGTATCCGGGAGAGGTAAAGGCAATGGTGGGGCATTTCAAG	313
Db	568	TATCTGGGAGAGGAAACAAATGTATCCGGGAGAGGTAAAGGCAATGGTGGGGCATTTCAAG	509
QY	314	CTAACATATCTGGTCGGGGACGAGCGTTGAGACAGAAATGATATTAACAACTTTGTGCAC	373
Db	508	CTAACATATCTGGTCGGGGACGAGCGTTGAGACAGAAATGATATTAACAACTTTGTGCAC	449
QY	374	CCCCACCTGTATCTGGCCCTCTTTTGAAGAGAGATGCAATTTGGCAGGCAAGAGAGTT	433
Db	448	CCCCACCTGTATCTGGCCCTCTTTTGAAGAGAGATGCAATTTGGAGGCAAGAGAGGTT	389
QY	434	CTGCCTCAGACACAGCTGTGACGAGATTTCCTGAAGTGGAGGATGATGTGATTAATGCTT	493
Db	388	CTGCCTCAGACACAGCTGTGACGAGATTTCCTGAAGTGGAGGATGATGTGATTAATGCTT	329
QY	494	CTGAGGAAGAGATGATATTCGATGCTTTGGATGATTTCTGATGACGACCTTGCAAGTATG	553
Db	328	CTGAGGAAGAGATGATATTCGATGCTTTGGATGATTTCTGATGACGACCTTGCAAGTATG	269
QY	554	ATTATGATCTCGGATGTGATCTCAAAAGAGCATGAGATCAAGAAAGCAGATTAATGGTGTCA	613
Db	268	ATTATGATCTCGGATGTGATCTCAAAAGAGCATGAGATCAAGAAAGCAGATTAATGGTGTCA	209
QY	614	AAAAGTCTTTTGGCAGCTTGATAGCTTTGTCATCGACGAGATTAATGAAACCAAGAGGC	673
Db	208	AAAAGTCTTTTGGCAGCTTGATAGCTTTGTCATCGACGAGATTAATGAAACCAAGAGGC	149
QY	674	AGTGGCATTTGTCAGTTGTCAAGACGACCTGTGTCCATCGATTGTTAATCTTGACACC	733
Db	148	AGTGGCATTTGTCAGTTGTCAAGACGACCTGTGTCCATCGATTGTTAATCTTGACACC	89
QY	734	CTCTACTAGCTATGC 749	
Db	88	CTCTACTAGCTATGC 73	

RESULT 2	606 bp	DNA	linear	GSS 28-JAN-2000
A0959657/c				
LOCUS	A0959657			
DEFINITION	LERM21IF LERE Arabidopsis thaliana genomic clone LERM21, genomic survey sequence.			

ACCESSION	A0959657	GI:6787358
VERSION	A0959657.1	
KEYWORDS	GSS.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta, eudicotyledons, core eudicots; Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsi-
1 (bases 1 to 606)
Buell C B, Yin Y, Dai G, Barnstead M, Bowman C, Treibach, T. et al. 2006

FEATURES

```
source
1..506
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="LANDSBERG ERECTA"
/db_xref="taxon:3702"
```

ORIGIN

Query Match	28.4%	Score 534	DB 28	Length 606
Best Local Similarity	100.0%	Pred. NO.	4.5e-118	
Matches 534	0	Mismatches	0	Indels 0
				Gaps 0

QY	216	AGCTTGGGGTGTGTGTCAGACGACAAGGAGAGGTAGCAACGTATCTGGGAGAGGAAAACAATGT	275
Db	606	AGCTTGGGGTGTGTGTCAGACGACAAGGAGAGGTAGCAACGTATCTGGGAGAGGAAAACAATGT	547
QY	276	ATCCGGGAGAGGTAAACGGCAATGGTCCGGGGCACTTCAAGCTAATCAATATCTGTCCGGGGACG	335
Db	546	ATCCGGGAGAGGTAAACGGCAATGGTCCGGGGCACTTCAAGCTAATCAATATCTGTCCGGGGACG	487
QY	336	AGCCTTAGGACGAAAGATGATTAACAACCTTGTGTGGACCCGCCCACTGTATCTGCACCCTCC	395
Db	486	AGCCTTAGGACGAAAGATGATTAACAACCTTGTGTGGACCCGCCCACTGTATCTGCACCCTCC	427
QY	396	TTTGGAAAGAGAGATGATAATTGGCAGGCAAGAGAGGTTCTGTCTCAGACAACAGCTGTGCA	455
Db	426	TTTGGAAAGAGAGATGATAATTGGCAGGCAAGAGAGGTTCTGTCTCAGACAACAGCTGTGCA	367
QY	456	CGAGTTTCTCTGACCTGAGAGATGATGTGTGATTAATGCTTCTGAGGAAAGAGATGATTTCCGA	515
Db	366	CGAGTTTCTCTGACCTGAGAGATGATGTGTGATTAATGCTTCTGAGGAAAGAGATGATTTCCGA	307
QY	516	TGCTTTGGATGATTTCTGTATGCAGACTTGCACAGATGATTAATGACTCCGATGTGATCA	575
Db	306	TGCTTTGGATGATTTCTGTATGCAGACTTGCACAGATGATTAATGACTCCGATGTGATCA	247
QY	576	AAAGAGCCATGATCAACGAAAGCAGAAATTAAGTGTTCAAAAGTTCCTTTGGCAGCTTGGGA	635
Db	246	AAAGAGCCATGATCAACGAAAGCAGAAATTAAGTGTTCAAAAGTTCCTTTGGCAGCTTGGGA	187
QY	636	TAGCTTGTGCATTCGACGAGATTAATGAAACCAAGAGCAATGGGATGTTCCAGCTTGCA	695
Db	186	TAGCTTGTGCATTCGACGAGATTAATGAAACCAAGAGCAATGGGATGTTCCAGCTTGCA	127
QY	696	GAAAGGACCTGTGTGCATTCGATTTGGTATTAACCTGCACCCCTTACTAGCTCATGC	749
Db	126	GAAAGGACCTGTGTGCATTCGATTTGGTATTAACCTGCACCCCTTACTAGCTCATGC	73

RESULT

AU236368 647 bp mRNA linear EST 01-APR-2002
 LOCUS AU236368 RAF1.4 Arabidopsis thaliana cDNA clone RAF1.4-93-K05 5',
 DEFINITION mRNA sequence.
 ACCESSION AU236368
 VERSION AU236368.1 GI:19875537
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 647)
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
 and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 Unpublished (2002)
 Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda FIC-1 vector (Carlucci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified pluescript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/planet/index_e.html) for further
 details.

FEATURES

source

Location/Qualifiers
 1..647
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF1.4-93-K05"
 /issue_type="root"
 /lab_host="DH10B"
 /clone_1ib="RAF1.4"
 /note="Site_1: BamHI; Site_2: SalI"

ORIGIN

Query Match 28.4%; Score 533; DB 9; Length 647;
 Best Local Similarity 99.5%; Pred. No. 8.1e-118;
 Matches 555; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
 1 ATAGATTCTAGGGCTGCTCCATGCTTAAGGAAAGAAAGCTTACGGGTGTTATAGGCTT 60
 92 ATGAGTTCTAGGGCTGCTCCATGCTTAAGGAAAGAAAGCTTACGGGTGTTATAGGCTT 151
 61 GAGGTGAACAGTTGGTTCAAGGTTTGGACGAGACGAGCTGCTTCTTCAAGAATGAT 120
 152 GAGGTGAACAGTTGGTTCAAGGTTTGGACGAGACGAGCTGCTTCTTCAAGAATGAT 211
 121 GAGGAGAGTGGAGGCTATTTCAGAGAGACAAAGAACAAACCAAGCACTTTGGA 180
 212 GGAGGAGAGTGGAGGCTATTTCAGAGAGACAAAGAACAAACCAAGCACTTTGGA 271
 181 AAAAAGTGGGTTCTCAGAAATTCGATCCCTAGAGCTTGGGTGCTCAGCAAGAGG 240
 272 AAAAAGTGGGTTCTCAGAAATTCGATCCCTAGAGCTTGGGTGCTCAGCAAGAGG 311
 241 AGAGTAGCAACGATCTGGGAGAGGAAACATGATCCGGGAGAGGTACGGCAATGCT 300
 332 AGAGTAGCAACGATCTGGGAGAGGAAACATGATCCGGGAGAGGTACGGCAATGCT 391
 301 CGGGGATTCACGATCAATATCTGGTCCGGGACGAGCTTGGACGAAAGTATGATAAC 360
 392 CGGGGATTCACGATCAATATCTGGTCCGGGACGAGCTTGGACGAAAGTATGATAAC 451

QY 361 AACTTTGTGGCAACCCCACTGATCTGCCTCTTTGAAAGAGATGAATTGGCAG 420
 DB 452 AACTTTGTGGCAACCCCACTGATCTGCCTCTTTGAAAGAGATGAATTGGCAG 511
 QY 421 GCAGAGAGAGGTTCTGCTCAGACACAGCTGTGACAGAGTTCTTCACTGGAGATGAT 480
 DB 512 GCAGAGAGAGGTTCTGCTCAGACACAGCTGTGACAGAGTTCTTCACTGGAGATGAT 571
 QY 481 GTGATTAATGCTTCTGAGAGAGATGATCCGATGCTTTGGATGATCTGATGACGAC 540
 DB 572 GTGATTAATGCTTCTGAGAGAGATGATCCGATGCTTTGGATGATCTGATGACGAC 629
 QY 541 CTTCGAAAGTATGATTTAT 558
 DB 630 CTTCGAAAGTATGATTTAT 647

RESULT 4
 A0964580/c 613 bp DNA linear GSS 28-JAN-2000
 LOCUS LERGX20TF LERG Arabidopsis thaliana genomic clone LERGX20, genomic
 DEFINITION survey sequence.
 ACCESSION A0964580
 VERSION A0964580.1 GI:6792281
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 613)
 Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Uterbach, T.,
 Feldblyum, T., Liang, F., Creasy, T. and Fraser, C.M.
 Genomic survey sequencing of Landsberg erecta ecotype of
 Arabidopsis thaliana and identification of sequence-based
 polymorphisms
 Unpublished (2000)
 Contact: Xiaoying Lin
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: atc@igr.org
 For additional information, see <http://www.tigr.org/cdb/at/at.html>
 Seg primer: TP
 Clase: shocgun.

FEATURES

source

Location/Qualifiers
 1..613
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="LANDSBERG ERECTA"
 /db_xref="taxon:3702"
 /clone="LERGX20"
 /clone_1ib="LERGX20"
 /note="Organ: Leaf; Vector: pUC19/JK; Total genomic DNA was
 sheared to 0.4-0.7 Kbp before ligation."

ORIGIN

Query Match 27.9%; Score 523.4; DB 28; Length 613;
 Best Local Similarity 99.8%; Pred. No. 1.7e-115;
 Matches 524; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 ATGAGTTCTAGGGCTGCTCCATGCTTAAGGAAAGAAAGCTTACGGGTGTTATAGGCTT 60
 DB 525 ATGAGTTCTAGGGCTGCTCCATGCTTAAGGAAAGAAAGCTTACGGGTGTTATAGGCTT 466
 QY 61 GAGGTGAACAGTTGGTTCAAGGTTTGGACGAGACGAGCTGCTTCTTCAAGAATGAT 120
 DB 465 GAGGTGAACAGTTGGTTCAAGGTTTGGACGAGACGAGCTGCTTCTTCAAGAATGAT 406
 QY 121 GAGGAGAGTGGAGGCTATTTCAGAGAGACAAAGAACAAACCAAGCACTTTGGA 180
 DB 405 GGAGGAGAGTGGAGGCTATTTCAGAGAGACAAAGAACAAACCAAGCACTTTGGA 346

QY 181 AAACTTGGTTCTCAGATTGCAATCTCTAGAGCTTGGGTGCTGACGACGAGG 240
 DB 345 AAACTTGGTTCTCAGATTGCAATCTCTAGAGCTTGGGTGCTGACGACGAGG 286
 QY 241 AAGGTAGCAACGATCTCTGGAGAGGAAACAATGATCCGGAGAGGTAAACGCAATGAT 300
 DB 285 AAGGTAGCAACGATCTCTGGAGAGGAAACAATGATCCGGAGAGGTAAACGCAATGAT 226
 QY 301 CGGGGCAATTCAGCTAACATATCTGGTCCGGGACGAGCGTTGACGAAAGTATGATTAAC 360
 DB 225 CGGGGCAATTCAGCTAACATATCTGGTCCGGGACGAGCGTTGACGAAAGTATGATTAAC 166
 QY 361 AACTTTGGGCAACCCCACTGTATCTGGCCTCTTTGAAAGAGATGAAATTTGGCAG 420
 DB 165 AACTTTGGGCAACCCCACTGTATCTGGCCTCTTTGAAAGAGATGAAATTTGGCAG 106
 QY 421 GGAAGAGAGGTTCTGCTGACGACAGCGTGTGACGAGATTTCTGACGTGAGAGTAT 480
 DB 105 GGAAGAGAGGTTCTGCTGACGACAGCGTGTGACGAGATTTCTGACGTGAGAGTAT 46
 QY 481 GTGATTAATGCTTCTGAGAGAGATGATTCGATGCTTTGGAT 525
 DB 45 GTGATTAATGCTTCTGAGAGAGATGATTCGATGCTTTGGAT 1

RESULT 5
 AVS66465/c 619 bp mRNA linear EST 07-SEP-2000
 LOCUS AVS66465 Arabidopsis thaliana green siliques Columbia Arabidopsis
 DERIVATION thaliana cDNA clone SQ244b06F 3', mRNA sequence.
 ACCESSION AVS66465
 VERSION AVS66465.1 GI:8737917
 SOURCE EST.
 ORGANISM Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 619)
 Asamizu, E., Nakamura, Y., Sato, S., and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 DNA Res. 7 (3), 175-180 (2000)
 JOURNAL MEDLINE
 PUBMED 20363093
 COMMENT 10907847
 Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1..619
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="SQ244b06F"
 /tissue_type="green siliques"
 /clone_1fb="Arabidopsis thaliana green siliques Columbia"
 /note="Vector: pBluescriptII SK-, Site_1: EcoRI, Site_2:
 XhoI"

FEATURES
 source
 ORIGIN
 Query Match 27.3%; Score 513; DB 9; Length 619;
 Best local Similarity 100.0%; Pred. No. 5.5e-113;
 Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 CAAGAGATGTTGTAAGAGAGCTGAGCAAGATCTCTGAGCAATCAGCAGCTGAATAC 1350
 DB 619 CAAGAGATGTTGTAAGAGAGCTGAGCAAGATCTCTGAGCAATCAGCAGCTGAATAC 560

QY 1351 TTTAAGAACAGCTCTCAAAACAGAACAGCAGCCCAAGTGTCTTGAGAAATCTCTGAA 1410
 DB 559 TTTAAGAACAGCTCTCAAAACAGAACAGCAGCCCAAGTGTCTTGAGAAATCTCTGAA 500
 QY 1411 ATTATGACGAGAAAGCTGCTGAGAACTGACAGAGATTAATCGATCTGACAGAGAACT 1470
 DB 499 ATTATGACGAGAAAGCTGCTGAGAACTGACAGAGATTAATCGATCTGACAGAGAACT 440
 QY 1471 AAGATGACGATGAAACAGAAACAGGAGAAAGATGATGACAGCAGAGTTTTCATGAT 1530
 DB 439 AAGATGACGATGAAACAGAAACAGGAGAAAGATGATGACAGCAGAGTTTTCATGAT 380
 QY 1531 TCAATCAACAGATTCATGAAAGAAAGAGACCCAAAGAGAGAAATTTGAGATGTTGCAG 1590
 DB 379 TCAATCAACAGATTCATGAAAGAAAGAGACCCAAAGAGAGAAATTTGAGATGTTGCAG 320
 QY 1591 CAGCAGAAACGTGCCAAGTTGTTGGCCAGACAGACAGCAAAATTAATCCTCTTGCAAT 1650
 DB 319 CAGCAGAAACGTGCCAAGTTGTTGGCCAGACAGACAGCAAAATTAATCCTCTTGCAAT 260
 QY 1651 GACGATTCGCCAAAGAGAGCTGAGAAAGTCAAGCTTCATGAGTTCAAGAGAAAGAG 1710
 DB 259 GACGATTCGCCAAAGAGAGCTGAGAAAGTCAAGCTTCATGAGTTCAAGAGAAAGAG 200
 QY 1711 ATGAGAGAGTTGTTGAGAGAGAGAGATGCTGATTAAGATCAAGAGAAAGATGGA 1770
 DB 139 ATGAGAGAGTTGTTGAGAGAGAGAGATGCTGATTAAGATCAAGAGAAAGATGGA 140
 QY 1771 GACATGAAGAGAGGATCAGAGAGATATT 1803
 DB 139 GACATGAAGAGAGGATCAGAGAGATATT 107

RESULT 6
 AQ959659 523 bp DNA linear GSS 28-JAN-2000
 LOCUS AQ959659 LEREM21TR LERE Arabidopsis thaliana genomic clone LEREM21, genomic
 DERIVATION survey sequence.
 ACCESSION AQ959659
 VERSION AQ959659.1 GI:6787360
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 523)
 Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T.,
 Feldblum, T., Liang, F., Creasy, T., and Fraser, C.M.
 Genomic survey sequencing of Landsberg erecta ecotype of
 Arabidopsis thaliana and identification of sequence-based
 polymorphisms
 unpublished (2000)
 Contact: Xiaoying Lin
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: at@cigr.org
 For additional information, see http://www.tigr.org/cdb/at/at.html
 Seq primer: TR
 Class: Shotgun.
 Location/Qualifiers
 1..523

FEATURES
 source
 ORIGIN
 Query Match 27.3%; Score 513; DB 9; Length 619;
 Best local Similarity 100.0%; Pred. No. 5.5e-113;
 Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 CAAGAGATGTTGTAAGAGAGCTGAGCAAGATCTCTGAGCAATCAGCAGCTGAATAC 1350
 DB 619 CAAGAGATGTTGTAAGAGAGCTGAGCAAGATCTCTGAGCAATCAGCAGCTGAATAC 560

```

FEATURES
source
location/Qualifiers
1. 644
/organism="Brassica napus"
/mol type="rRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"

```

RESULT	8
CDB15300	
LOCUS	
DEFINITION	708 bp mRNA linear EST 10-JUL-2003 BN15.02500P020214 BN15 Brassica napus cDNA clone BN15025002, mRNA sequence.
ACCESSION	CDB15300
VERSION	CDB15300.1 GI:32497240
KEYWORDS	EST.
SOURCE	Brassica napus (rape)
ORGANISM	Brassica napus
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica. 1 (bases 1 to 708) Genoplatne . Genoplatne, a major partnership french program in plant genomics Unpublished (2003) Contact : Genoplante Genoplante
TITLE	
JOURNAL	
COMMENT	

93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' (<http://www.genoplante.com>
 and <http://genoplante-info.infobiogen.fr>).

FEATURES

source

1..708
 /organism="Brassica napus"
 /mol_type="mRNA"
 /cultivar="Jec neuf"
 /db_xref="taxon:3708"
 /clone="BN15025002"
 /tissue_type="seed"
 /clone_lib="BN15"

ORIGIN

Query Match 23.0%; Score 431.2; DB 14; Length 708;
 Best Local Similarity 78.1%; Pred. No. 3,6e-93;
 Matches 575; Conservative 0; Mismatches 118; Indels 43; Gaps 3;
 1057 GGGATGAGTGTCTGATGTTGAGAGACGTCACCTGCTATTGAGGCGGACGCTTC 1116
 Db 1 GGGATTAGTGTCTGATCTTTGAGAGTACCGCCACTGGCTACTTTGAGGCTGACCGCTTA 60
 QY 1117 CAC-CGGAGTTAGCTGAGATGGGCTTACATGATTTGCTGGGGTTCAGAGGCGCATAT 1175
 Db 61 CACAGGGGACCTTGTCTCAAAAGGGGTTGATGAGAGCGCTGGAGCCGTTGCCGAGATG 120
 QY 1176 GTTTTCGAGAGTGTGGCCACTGTATGGCTTCCTTCAACAGAGCAAGATCTGACAT 1235
 Db 121 GTTTTCGAGAGTGTGGCCACTGTATGGCTTCCTTCAACAGAGCAAGATCTGACAT 180
 QY 1236 ATTCATCAACACTCTCA--AGGCAAAACAGGCTGAAATTCGAGTTGAATCATACCA 1292
 Db 181 ATTCATCAACACTCTCAAGTAGGAGAAACAGGCTGACATGAGTGAATCTGATCCA 240
 QY 1293 AGAGATGTTGTAAAGAGCTGAGGCGATCTCTGAGAGCAATCAGAGCTGAACTT 1352
 Db 241 GAGATAGTGTGAGAGAGCTGAGGCGATCTCTGAGAGCAATCAGAGCTGAACTT 300
 QY 1353 TAAGAACAGACTCTCAAAAGAGAGCAAGCAGCCAGGCTCTGAGGATCTCTGGAAT 1412
 Db 301 CAGAGACAGCTGACGAAAGAGAGCAAGCAGCCAGGCTCTGAGGATCTCTGGAAT 360
 QY 1413 TATGAGGAGAGAGCTGCTAGAACTGCAAGAGATATCGATGCTGAGACAGAACTAA 1472
 Db 361 TTTCAGCGAGAGAGCTGCTAGAACTGCAAGAGATATCGATGCTGAGACAGAACTAA 420
 QY 1473 GATGACAGATGAGAGAGAGAGAGAGATGATGACACAGACAGGTTTTCATGATTC 1532
 Db 421 GATGACAGATGAGAGAGAGAGAGAGATGATGACACAGACAGGTTTTCATGATTC 480
 QY 1533 AATCAACAGATCCATGAAAGAGAGAGAGAGAGAGATTTCCAGATGTTGACGA 1592
 Db 481 CATTAAGCAGATCATGAGAGAGAGAGAGAGAGAGAGATTTCCAGATGTTGACGA 540
 QY 1593 GCAAGAACGTGCAAGGTTGTTGGCCAGACAGACAGACAGAACTTAATCCTCTAGCAATGA 1652
 Db 541 GCATGACGTTCTTAAGTTGTT----- 562
 QY 1653 CGATGCCGAAAGAT 1712
 Db 563 -GATTTCCGAAAGAT 621
 QY 1713 GAGAGAGTTGTGAT 1772
 Db 622 GAAAGAGTTTGTGAT 681
 QY 1773 CATGAGAGAGAGAGAT 1788
 Db 682 CTGAAAGAAAAATAT 697

RESULT 9

CE517786

LOCUS

DEFINITION

CE517786 849 bp mRNA linear EST 09-SEP-2003
 CAP0005_IF_E12 Vitis vinifera cv. Cabernet Sauvignon (Clone 8)
 Petiole - CAP Vitis vinifera cDNA clone CAP0005_IF_E12 5', mRNA
 sequence.

ACCESSION

CE517786

KEYWORDS

SOURCE

ORGANISM

EST

Vitis vinifera

Vitis vinifera

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 849)
 Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and
 Cook, D.
 Expressed sequence tags from Vitis vinifera 'Cabernet Sauvignon'
 berries at various developmental stages
 Unpublished (2003)
 Contact: Douglas Cook, PhD
 CAES Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 Seq primer: ACGTACCGACATATGCC.
 Location/Qualifiers
 1..849
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAP0005_IF_E12"
 /sex="Hermaphrodite"
 /dev_stage="Onset of Veraison (berry softening)"
 /lab_host="DMSalpa"
 /clone_lib="Vitis vinifera cv. Cabernet Sauvignon (Clone
 8) Petiole - CAP"
 /note="Organ: Petiole; Vector: pDNR; Site: 1: SfiI; Site: 2:
 SfiI; CAP is a cDNA library of Vitis vinifera cv.
 'Cabernet Sauvignon' Clone 8 petioles. Samples were
 collected on July 10, 2002 from plants on the onset of
 Veraison (berry softening). Sampled vines were located at
 the University of California, Davis, Experimental
 Vineyard. cDNAs were made by oligo-dT priming and
 directionally cloned. 5' and 3' adaptors were used in
 cloning as follows:
 5'-AAGCACTGATCAACGACAGAGTGGCCATTAGGCGCGG-3' and
 5'-ATTCTAGAGCCGAGCGGCCACATAG-dT(30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 20.8%; Score 390; DB 14; Length 849;
 Best Local Similarity 68.5%; Pred. No. 3,6e-83;
 Matches 569; Conservative 0; Mismatches 256; Indels 6; Gaps 2;
 469 GTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528
 Db 19 GAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 78
 QY 529 TCTGATGACGACCTTGCAAGTATGATGATGATGATGATGATGATGATGATGAT 588
 Db 79 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 138
 QY 589 TCAGGAAGCAGAAATAGTGTCAAAAGTCTTTGGAGAGCTTGTGATGCTTGTGATC 648
 Db 139 ACAGGAT 198
 QY 649 GAGCAGATTAATGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 708

```

Db 199 GAGGAATTAACGACACGACGAGTGGCATTTGCTTCATGCCAAGGTGACCTGGT 258
QY 709 GCGATGATTTGATTA---ACCTGCAACCTCTACTAGCTATCCGAGAACAAAGAGCT 765
Db 259 GCGATTGACCTGATACAGGGGTCTCCAACTTATCAACATGCCAAACAAAGAGATCA 318
QY 766 AGCGAGATTAACTCCATAGAGATTGGCTGGAAGTTTAAAGAAAGATCTACAGATGAGA 825
Db 319 MAGAGGTCAAACTTCATCGAGAGCTTGCAAACTTCTGATGAGAGATTGTACAGAGG 378
QY 826 GGGCATCTGTCTATCTCTGTGTGATTTATGGGCAAGAAAGGTTTGGTGAAGAT 885
Db 379 GGTACTTCTGTGTACAGCTGAGAGGACATTTGAAAGTGGAAGGTCATATATGAGTCG 438
QY 886 GAAAGAGATTAAATTTCTCTGCTCCCAATGTCTCTCATGATATCTAAGCTGAT 945
Db 439 GTTAAAGATCATAGATTGTTTGCCCTCCAAATGGTTATCATTAATGAACACTAGACTTGA 498
QY 946 AAGGAGATTAAGATTAAGTGGCTCGGATGGGCAACCAAGAGCTGTAATCTTGAC 1005
Db 499 CAGGATCTGATGAGAGTGGATTTGGGATGGGATTAAGAGCTTCTTGATTTATTCAGC 558
QY 1006 AAGTATAGGCTCTTAGAGACGCCATTCCTATGTCCACAGGCGCATCTGGAGTAT 1065
Db 559 TCATATGCTGACATGAAGGCTCGACATTCCTATGTCTCAAGAGGCAAGAGGATAGT 618
QY 1066 GTTCTGATGTTTGAAGACAGTGCACATGCTATTTGAAGGCCAGCCCTCCACGGGAG 1125
Db 619 GTTTTATTTTAAAGCTGGGCAATTTGTTACTTAAAGGCTGAACGCCCTTCAATAGCAC 678
QY 1126 TTAGCTGAGATGGGGTTAGATAGATTTGCTGGGGGTGAGAAAGC---CAGTATGTTTCT 1182
Db 679 TTGCTGAGACAGGACAGATTAAGGAAGCATGGATGCTGTCTGCTCTTCTATCTCT 738
QY 1183 GGAAGTGTTCGCACTGATGAGCTTCTCTTCAACGAGCAAGATCTGAATATTCAT 1242
Db 739 GGAATCNGCGCAGCTTTATGATCATGGCTATAAAGAAATTTGAACATTCATCAAC 798
QY 1243 CAACACTCTTAAGGCAAAACAGGCTGAATTTGAGTTGAATCATTCACA 1293
Db 799 CAACATCTCAAGAAATCAAAACGATTTGAGATGATCATATCA 849

```

```

RESULT 10
LOCUS AV525508 380 bp mRNA linear EST 01-SEP-2000
DEFINITION AV525508 Arabidopsis thaliana aboveground organs two to six-week
ACCESSION AV525508
VERSION AV525508.1 GI:8685036
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 380)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20361093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source 1..380
/organism="Arabidopsis thaliana"

```

```

/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="APD25d02R"
/issue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_1b="Arabidopsis thaliana aboveground organs two to
six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

```

ORIGIN

```

Query Match 20.1%; Score 378.4; DB 9; Length 380;
Best Local Similarity 99.7%; Pred. No. 1,6e-80;
Matches 379; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 276 ATCCGGAGAGGTACCGCAATGTTGGGCAATTCAGCTAACATATCTGTGGGGAGG 335
Db 1 ATCCGGAGAGGTACCGCAATGTTGGGCAATTCAGCTAACATATCTGTGGGGAGG 60
QY 336 AGCGTTAGCAGAAATATGATTAACAATTTGTCGACCCCACTGTATCTGACCTCC 395
Db 61 AGCGTTAGCAGAAATATGATTAACAATTTGTCGACCCCACTGTATCTGACCTCC 120
QY 396 TTGGAAGAGATGAAATTTGGCAGGCAAGAGAGTTCTGTCAACACAGCTGTGA 455
Db 121 TTGGAAGAGATGAAATTTGGCAGGCAAGAGAGTTCTGTCAACACAGCTGTGA 180
QY 456 GGAAGTTCTGACGTGAGAGATGATGATTAATGCTTCTGAGAAAGAAATATCCGA 515
Db 181 GGAAGTTCTGACGTGAGAGATGATGATTAATGCTTCTGAGAAAGAAATATCCGA 240
QY 516 TGCTTGAATGATTCGATGAGACGCTTCAAGTATGATTAATGATCGAGTGTAGTGA 575
Db 241 TGCTTGAATGATTCGATGAGACGCTTCAAGTATGATTAATGATCGAGTGTAGTGA 300
QY 576 AAAGACCAATGATCAGAAAGAGATGATGATTAATGATTAATGATCGAGTGTGA 635
Db 301 AAAGACCAATGATCAGAAAGAGATGATGATTAATGATTAATGATCGAGTGTGA 360
QY 636 TAGCTTGTGATGAGCAGA 655
Db 361 TAGCTTGTGATGAGCAGA 380

```

```

RESULT 11
LOCUS CF517260/c 808 bp mRNA linear EST 09-SEP-2003
DEFINITION CAP0005 IR_E12 Vitis vinifera cv. cabernet sauvignon (Clone 8)
petiole - CAP Vitis vinifera cDNA clone CAP0005_IR_E12 3', mRNA
sequence.
ACCESSION CF517260
VERSION CF517260.1 GI:34549028
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
REFERENCE 1 (bases 1 to 808)
AUTHORS Goes da Silva, F., Iandolo, A., Lam, H., Baek, J., Jones, K. and
Cook, D.
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: GCCAAGCAATGCTTAG.

```

Location/Qualifiers

ORIGIN

Matches 538; Conservative 0; Mismatches 247; Indels 6; Gaps 2.

808 AAGAAGACCAAGTGGTTTAAGTCATTTTTTGAGATATTGGATAGCTNATCATTGAGGAA 749

100

[illegible]

772 GTTACGCTCCATAGAGAATTGGCTGAAGTTTTAGAAAAGGATCTACAGATGAGAGGCCA 831

832 TCTGTCATTCCCTTGTGTGAGATTATGGCAGTGAAGGTTTGGGTGAGGATGAAGA 891

9E1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 448 GCTGATGAGAAGTGGATTGGGATGGGAATCAAGAGCTTCTTGATTATTTCAGCTCATAT 389

Db 388 GCTGCAGTGAAGGCTCGACATTCTATGTTCCACACAAGGCGACAGAGGAATGAGTGTTTTG 329

[illegible]

α

1189 GTTGGCACTGTATGGCTTCCCTTGCAACGAAGCAAGATCTGGACATATTCAATCAACAC 124

1249 TCTCAGGCAAAACACAGGCTGAATTTCGATTGAATCATACCAAGAGATGTTGTAAG 1308

1309 GAGCTGAGCGAGATCTCTGAGGACAATCAGACGCTGAACTTAAGAACAGCTCTCA 1368

[illegible]

ACCESSION	AI999551
VERSION	AI999551.1
	GI:5846456

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Iacineophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Tan, R., Rose, M., Warren, B., Itoh, B., Kasbury, K., Borlino, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,

Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.

```
source      1. .548
/organism="Arabidopsis thaliana"
```

```

/aev stages="4 - /weaks
/clone_lib="A. thaliana, Columbia Col-0, rosette-3"

```

watered with fertilizer. CDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was

Query Match	19.2%;	Score 359.8;	DB 9;	Length 548;
Best Local Similarity	97.0%;	Pred. No. 6e-76;		

10

Db 548 GCACACCACAGGTTTTCATGATTCATCAATCAGATCCATGAAGAAGAGACGCAAG 489
 QY 1567 GAGAGAAATTCGATGTTGCACAGCAGAAAGTTCGAAGTTGTGGCCAGCAG 1626
 Db 488 GAGAGAAATTCGATGTTGCACAGCAGAAAGTTCGAAGTTGTGGCCAGCAG 429
 QY 1627 CAGAACTTAATCCCTCTAGCAATGACATGTCGAAAGAGCTGAGAAAGTGTCAAGC 1686
 Db 428 CAGAACTTAATCCCTCTAGCAATGACATGTCGAAAGAGCTGAGAAAGTGTCAAGC 369
 QY 1687 TTGATGAGTTTCAAGAAAGATGAGAGATTTGTGAAAGAGGGGAGATGCTGATA 1746
 Db 368 TTGATGAGTTTCAAGAAAGATGAGAGATTTGTGAAAGAGGGGAGATGCTGATA 309
 QY 1747 AAAGATCAAGAAAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGATGAT 1806
 Db 308 AAAGATCAAGAAAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGATGAT 249
 QY 1807 CTGAGAAAGATTTGATGAGCTTTGGAACAGCTCATGTAAGAGATGAGCTTCAAT 1866
 Db 248 CTGAGAAAGATTTGATGAGCTTTGGAACAGCTCATGTAAGAGATGAGCTTCAAT 189
 QY 1867 GAAGATGATTTGA 1878
 Db 188 GAAGATGATTTGA 177

RESULT 13
 CA18697
 LOCUS 785 bp mRNA linear EST 09-DEC-2002
 DEFINITION CA12E13041IF G03 Cabernet Sauvignon Leaf - CA12E1 Vitis vinifera
 ACCESSION CA18697
 VERSION CA18697.1 GI:26265634
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 785)
 Goes da Silva, F., Lim, H., Iandolo, A., Baek, J., Jones, K.,
 Walker, M.A. and Cook, D.R.
 Transcriptional responses of Vitis vinifera to infection by the
 bacterial pathogen Xylella fastidiosa
 Unpublished (2003)
 Contact: Doug Cook
 CAES Genome Facility
 UC Davis Department of Plant Pathology
 1 Shields Ave., Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 Seq primer: GTTATCATGTCACGCTAC.
 Location/Qualifiers
 1..785

FEATURES
 source
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultiivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CA12E13041IF G03"
 /sex="hermaphrodite"
 /dev_stage="mid-season leaf material"
 /lab_host="DH5alpha"
 /clone_lib="Cabernet Sauvignon Leaf - CA12E1"
 /note="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2:
 SfiI; CA12E1 is a cDNA library of Cabernet Sauvignon
 leaves. The leaves were collected on July 25, 2001, in
 Napa Valley, California, and represent leaves in
 mid-season development. These leaves were verified to be
 infected with the bacterial pathogen, Xylella fastidiosa,
 based on a diagnostic assay using PCR and Xylella-specific
 primer pairs. The plants were asymptomatic at the time of

collection, but later developed symptoms. cDNAs were made
 by oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AGCAGTGGTATCAACGACAGAGGAGGATTCAGCGCGG-3' and
 5'-ATTCTTAGAGCCGAGCGGCGGATG-3' (30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
 Query Match 19.1%; Score 358.2; DB 14; Length 785;
 Best Local Similarity 68.4%; Pred. No. 1.7e-75;
 Matches 527; Conservative 0; Mismatches 238; Indels 6; Gaps 2;

QY 495 TGAGAAAGAAATGATTCGATGCTTTGATGATTTCTGATGACGACCTTGCAAGTATGA 554
 Db 15 TGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74
 QY 555 TTATGACTCGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614
 Db 75 TTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 134
 QY 615 AAAGTTCTTTGAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 674
 Db 135 GTGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 194
 QY 675 GTGCAATGTCAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
 Db 195 GTGCAATGTCAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 254
 QY 732 CCCTCTACTAGCTCATGAGGAGCAAAAGAGCTAGGAGATTAAGCTCATAGAAATT 791
 Db 255 ACCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 314
 QY 792 GCGTGAAGTTTGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 851
 Db 315 TGCAAGAACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 374
 QY 852 GATTATGAGGAGTGAAGGTTTGGGTGAGTGAAGAAAGATTAAGAAATTGTGCGC 911
 Db 375 GGCATTTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 434
 QY 912 TCCATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 971
 Db 435 TCCAAATGTTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
 QY 972 CATGGCAACCAAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1031
 Db 495 GATGGGAAATCAAGAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 554
 QY 1032 TTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
 Db 555 TTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614
 QY 1092 TGGCTATTTGAGAGCGCAAGCTGCTCAACCGAGGATGATGATGATGATGATGAT 1151
 Db 615 TGGTACTTATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 674
 QY 1152 TGGCTGGGGTCAAGACG--CAGTATGTTTCTGAGAGTGTTCGCCAATGATGCTT 1208
 Db 675 AGCATGAGATGCTGCTGCTGCTTCTATCTGAGAGAAAGCGCACCTTATGATGATA 734
 QY 1209 CCTTGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
 Db 735 CATGGCTATTAAGAAAGATTTGACTATTAACCAACATTTCTCAAGAAA 785

RESULT 14
 BH535855
 LOCUS 697 bp DNA linear GSS 14-DEC-2001
 DEFINITION BOHRK13IF BOHK Brassica oleracea genomic clone BOHRK13, genomic
 survey sequence.
 ACCESSION BH535855
 VERSION BH535855.1 GI:17767651

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 19:34:49 ; Search time 94.7349 Seconds
(without alignments)
11001.196 Million cell updates/sec

Title: US-10-030-829-2
Perfect score: 1878
Sequence: 1 atgagcttcaggcgtgtcc.....ttccaatgaagatgatga 1878

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents NA:*
1: /cgn2_6/pcodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/pcodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/pcodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/pcodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/pcodata/2/ina/PCUTS_COMB.seq:*
6: /cgn2_6/pcodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	87.8	4.7	7218	1	US-08-232-463-14
2	75	4.0	272	4	US-09-313-294A-121
3	74.4	4.0	274	4	US-09-313-294A-463
4	58.8	3.1	3489	2	US-08-728-123A-1
5	58.8	3.1	3489	4	US-09-298-568-1
6	58.8	3.1	3489	4	US-09-410-399-1
C 7	58.8	3.1	32207	2	US-08-770-379-20
C 8	58.8	3.1	32207	3	US-08-757-669A-20
C 9	58.8	3.1	32207	4	US-09-230-371A-20
C 10	53.2	2.8	16442	3	US-08-781-891-208
C 11	53.2	2.8	16442	4	US-09-618-166-208
C 12	49.8	2.7	51259	3	US-08-781-891-209
C 13	49.8	2.7	51259	4	US-09-618-166-209
14	47.4	2.5	1276	3	US-09-411-812A-2
15	47.4	2.5	1276	4	US-09-590-113-2
16	47.4	2.5	1276	4	US-09-976-594-268
17	47.4	2.5	1276	4	US-09-976-594-268
18	46.6	2.5	2338	1	US-08-425-069-1
19	46.6	2.5	2338	2	US-08-425-069-1
20	45.4	2.4	340	1	US-08-182-175A-104
21	45.4	2.4	340	5	PCT-US92-06412-104
22	43	2.3	1926	4	US-09-249-585A-2
23	43	2.3	1926	4	US-09-410-399-3
24	43	2.3	2580	3	US-09-050-863-2
25	43	2.3	2580	4	US-09-359-081-2
C 26	43	2.3	5452	2	US-09-130-114-1
C 27	43	2.3	8705	4	US-09-647-344A-14

28	43	2.3	9600	3	US-08-910-647-1	Sequence 1, Appl
29	43	2.3	9600	4	US-09-620-925-1	Sequence 1, Appl
30	43	2.3	10596	1	US-07-884-811-15	Sequence 15, Appl
31	43	2.3	10596	1	US-07-885-971-15	Sequence 15, Appl
32	43	2.3	10596	1	US-08-087-783A-15	Sequence 15, Appl
33	43	2.3	10596	1	US-08-194-088B-15	Sequence 15, Appl
34	43	2.3	10596	2	US-08-194-087-15	Sequence 15, Appl
35	43	2.3	10596	5	PCT-US93-04648-15	Sequence 15, Appl
36	43	2.3	16080	4	US-09-724-565A-48	Sequence 48, Appl
C 37	42.2	2.2	1251	4	US-09-828-000-1	Sequence 1, Appl
C 38	42.2	2.2	1270	4	US-08-387-805-1	Sequence 1, Appl
39	42.2	2.2	1958	4	US-09-702-327-3	Sequence 3, Appl
40	41.2	2.2	2277	1	US-08-676-967-2	Sequence 2, Appl
41	41.2	2.2	2277	1	US-08-676-974-2	Sequence 2, Appl
42	41.2	2.2	2277	2	US-09-098-487-2	Sequence 2, Appl
43	41.2	2.2	8920	2	US-08-446-855A-1	Sequence 1, Appl
44	41.2	2.2	8920	3	US-09-150-741-1	Sequence 1, Appl
45	41	2.2	966	2	US-08-766-738-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ. ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PT9pct-F1s
; US-08-232-463-14
Query Match 4.7%; Score 87.8; DB 1; Length 7218;

Best Local Similarity 4.9%; Pred. No. 2,7e-15;
Matches 20; Conservative 250; Mismatches 137; Indels 0; Gaps 0;

QY 1388 AGGTGCTTGAAGATCTTGAAATTTATGACGAGAAAGCTGGTGAAGTGCAGAGATA 1447
DB 1465 AAGTATTGAAGATGAGAAATTTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1406
QY 1448 ATCGATCGTGACAGACAGAACTAAGATGACATGACGAGAAAGAGATGATG 1507
DB 1405 RRR 1346
QY 1508 CACACGACAGTTTTCATGATTCATCAATCAACAGATTCATGAAAGACGCAAG 1567
DB 1345 RRR 1286
QY 1568 AGGAAATTTGAGATTTTGCAGACAGAGAGTCCCAAGTTGTTGGCCAGACAGC 1627
DB 1285 RRR 1226
QY 1628 AGAATTAATCCCTAGCAATGACGATTCGCAAGAGAGCTGAGAGATGTCAAGT 1687
DB 1225 RRR 1166
QY 1688 TCATCGAGTTTCAAGAGAAAGATGAGAGATTTTGCAGAGAGAGAGATCTGATA 1747
DB 1165 RRR 1106
QY 1748 AAGATCAAGAGAGAGATGAGAGATGAGAGAGAGATCAAG 1794
DB 1105 RRR 1059

RESULT 2

US-09-313-294A-121
Sequence 121, Application US/09313294A

Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laljudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 121
LENGTH: 272
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6476212 700548569H1
NAME/KEY: unsure
LOCATION: 14, 32, 61, 127, 162, 192
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-121

Query Match

Best Local Similarity 4.0%; Score 75; DB 4; Length 272;
Matches 148; Conservative 0; Mismatches 106; Indels 3; Gaps 1;

QY 1009 TATGAGGCTTTAGAGACGCGCTTCTATGTCACAGAGGCGCATGTCGATGATGTT 1068
DB 2 TATGAAGCAAGTAAACAGCTCATGCTATNGTCCCTCTGGGACCGTGATGAGCGTN 61
QY 1069 CTGATGTTTGAAGAGAGTGCCTGCTATTTTGAAGCCGAAAGCCTCCACCGGAGTTA 1128
DB 62 TTAATATTGAAGAGCTCAGCTGCGGCTCATGAGAGCTGAGAGCTGATTAACCTTT 121
QY 1129 GCTGAGATGGGGTTAGATAGATTCCTGG--GCTCAGAGGCGAGTATGTTTCTGGA 1185
DB 122 GTTAANCAAGGTACAGACAGAAATTCATGCACTACGCAAGGTTGATTTGTGCTGTT 181

QY 1186 GGTGTCGCCAAGTGTATGCTTCTTGCACAGAACAGATCTGACATATTCATCA 1245
DB 182 GGGAAAGGCGNCCCTAATATGTTTCTTAGCAACAAAGAGATGACCATTTAAAG 241
QY 1246 CACTCTCAAGCAAAAC 1262
DB 242 CATGCCATGGTAAAGC 258

RESULT 3

US-09-313-294A-463
Sequence 463, Application US/09313294A

Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laljudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 463
LENGTH: 274
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6476212 700549278H1
NAME/KEY: unsure
LOCATION: 15, 53
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-463

Query Match 4.0%; Score 74.4; DB 4; Length 274;
Best Local Similarity 60.5%; Pred. No. 3.6e-12;
Matches 156; Conservative 0; Mismatches 98; Indels 4; Gaps 2;

QY 1009 TATGAGGCTTTAGAGACGCGCATTCATGATGCTTCCACAGGCGCATGTCGATGATGTT 1067
DB 2 TATGAAGCAAGTAAACAGCTCATGCTATNGTCCCTCTGGGACCGTGATGAGCGTN 61
QY 1068 TCTGATGTTTGAAGAGAGTGCCTGCTATTTGAGAGCGCGACCTCCACCGGAGTT 1127
DB 62 GTTAATATTGAAGAGCTCAGCTGCGGCTACATGAGAGCTGACATTAACACTTT 121
QY 1128 AGCTGAGATGGGTTAGATAGAAATTCCTGG--GCTCAGAGGCGAGATGTTTCTGG 1184
DB 122 TGTTAATCAAGGTACAGACAGAAATTCAGCACTACGCAAGGTTGATTTGTCCCTGG 181
QY 1185 AGGTGTTGCGCAACTGATATGCTTCTTCCAAAGAGAGATCTGACATATTCATCA 1244
DB 182 TGGGAAAGCACTATATGTTTCTTAGCAACAAAGAGATGAGGCAATTAACAA 241
QY 1245 ACACTCTCAAGCAAAAC 1262
DB 242 GCATTCGCAAGTAAAGC 259

RESULT 4

US-08-728-323A-1
Sequence 1, Application US/08728323A

Patent No. 5948676
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelmann, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: Immediate Early Protein From Kaposi's
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

```

; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,323A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52268/JPM/MSC/SKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3489 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3489
; US-08-728-323A-1

Query Match      3.1%; Score 58.8; DB 2; Length 3489;
Best Local Similarity 43.4%; Pred. No. 6.7e-07;
Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

QY 1218 GAAGCAGATCTGACATATTCATCACTCTCAAGGCAAAACAAGCTGAATTGCA 1277
DB 1800 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 1859
QY 1278 GTTGAATCATCAACAAGATGTTGTTAAAGAGCTGAGCAGATCTTGAAGCAATCA 1337
DB 1860 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 1919
QY 1338 GCAGCTGAATCTTTAAGAAAGCTCTCAAAACAAGCAAGCAGCCAAAGTGTGTA 1397
DB 1920 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 1979
QY 1398 GGAATCTCTGAAATTTATGACGAGAAAGCTGCTAGAACTGCAAGGATATCGATCGT 1457
DB 1980 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 2039
QY 1458 GAGACAGAACTTAAGATGACATGAACAAGAGGAGAGATGATGACACAGCAG 1517
DB 2040 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 2099
QY 1518 GTTTTCATGATTCATCAACAGATCCATGAAGAAAGAGCAAGCAAGGAGAAATTT 1577
DB 2100 GCAGCAGCAGATGACAGCAGAGATGAGCAGCAGCAGAGATGACAGCAGCA 2159
QY 1578 CGAGATTTGTCAGCAGCAGAAAGTGCCTGTTGTTGTCAGCAGCAGCAGAACTTAA 1637
DB 2160 GATGAGCAGCAGCAGAGATGAGCAGCAGCAGCAGATGACAGCAGAGATGAGCA 2219
QY 1638 TCCCTTACAGATGACGATTTGCCAAAGAGAGCTGAGGAAGTGTCAAAGCTTCAATCGATT 1697
DB 2220 GCAGCAGCAGCAGAGATGAGCAGCAGCAGCAGATGAGCAAGAGCAGAGAGAGCAGGA 2279

```

```

QY 1698 TCAGAAGAAAGATGAGAGAGATTGTGGAAGAGGAGATGCTGATTAAGATCAAGA 1757
DB 2280 GCAGCAGAGAGCAGCAGAGATGAGCAGAGATGAGCAGAGAGAGATGAGCAGCA 2339
QY 1758 GAAAGATGAGAAAGATGAGAAAGAGAGATGACAGAGATTTGATCTGAGAAAGA 1817
DB 2340 GAGGAGTTAGAGCAGCAGAGATGAGCAGAGATGAGCAGAGAGATGAGAGAGA 2399
QY 1818 ATTTGATGAGCTTGAACAG 1839
DB 2400 GCAGCAGAGATGAGAGAGAG 2421

RESULT 5
US-09-298-568-1
; Sequence 1, Application US/09298568
; Patent No. 6322792
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballester, Mary E.
; APPLICANT: Kaye, Kenneth M.
; TITLE OF INVENTION: RHADINO VIRUS DNA ACTS IN TRANS ON A UNIT OF RHADINO
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/298,568
; CURRENT FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,422
; EARLIER FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
; US-09-298-568-1

Query Match      3.1%; Score 58.8; DB 4; Length 3489;
Best Local Similarity 43.4%; Pred. No. 6.7e-07;
Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

QY 1218 GAAGCAGATCTGACATATTCATCACTCTCAAGGCAAAACAAGCTGAATTGCA 1277
DB 1800 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 1859
QY 1278 GTTGAATCATCAACAAGATGTTGTTAAAGAGCTGAGCAGATCTTGAAGCAATCA 1337
DB 1860 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 1919
QY 1338 GCAGCTGAATCTTTAAGAAAGCTCTCAAAACAAGCAAGCAGCCAAAGTGTGTA 1397
DB 1920 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 1979
QY 1398 GGAATCTCTGAAATTTATGACGAGAAAGCTGCTAGAACTGCAAGGATATCGATCGT 1457
DB 1980 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 2039
QY 1458 GAGACAGAACTTAAGATGACATGAACAAGAGGAGAGATGATGACACAGCAG 1517
DB 2040 GCAGCAGATGACAGCAGAGATGAGCAGCAGCAGATGACAGCAGAGATGAGCA 2099
QY 1518 GTTTTCATGATTCATCAACAGATCCATGAAGAAAGAGCAAGCAAGGAGAAATTT 1577
DB 2100 GCAGCAGCAGATGACAGCAGAGATGAGCAGCAGCAGAGATGACAGCAGCA 2159
QY 1578 CGAGATTTGTCAGCAGCAGAAAGTGCCTGTTGTTGTCAGCAGCAGCAGAACTTAA 1637
DB 2160 GATGAGCAGCAGCAGAGATGAGCAGCAGCAGCAGATGACAGCAGAGATGAGCA 2219
QY 1638 TCCCTTACAGATGACGATTTGCCAAAGAGAGAGCTGAGGAAGTGTCAAAGCTTCAATCGATT 1697
DB 2220 GCAGCAGCAGCAGAGATGAGCAGCAGCAGCAGATGAGCAAGAGCAGAGAGAGCAGGA 2279
QY 1698 TCAGAAGAAAGATGAGAGAGATTGTGGAAGAGGAGATGCTGATTAAGATCAAGA 1757

```

```

Db      2280 GCAGCAGGAGGAGCAGAGGAGTTAGAGCAGCAGAGGATTAAGAGTATAGCA 2339
Qy      1758 GAAGAAGATGGAAGACATGAAGAAGAGGCATCAAGAGATATTTGATCTGAGAAAGA 1817
Db      2340 GCAGAGATTAGAGCAGAGAGAGGATTAGAGAGCAGCAGAGATTAAGAGAGCA 2399
Qy      1818 ATTGATGAGGCTTTGGAACAG 1839
Db      2400 GGAGCAGAGATTAGAGAGAGCAG 2421

```

RESULT 6

```

US-09-410-399-1
; Sequence 1, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Collier, Murray A.
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-399-1

```

Query Match 3.1%; Score 58.8; DB 4; Length 3489;

Best Local Similarity 43.4%; Pred. No. 6.7e-07; Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

```

Qy      1218 GAAGCAGATCTGACATATTCATCACTCTCAAGGCAAAACAGGCTGAAATTCGA 1277
Db      1800 GCAGCAGATGACAGCAGCAGATGAGCAGCAGAGATGACAGCAGAGATGAGCA 1859
Qy      1278 GTTGAATCATACCAAGATGTTGTAAAGAGCTGAGGAGATCTTGAGGACATCA 1337
Db      1860 GCAGCAGATGACAGCAGCAGATGAGCAGCAGAGATGACAGCAGAGATGAGCA 1919
Qy      1338 GCAGCTGAATCTTTAAGAACAAAGCTCTCAAAACAGAAACAGCCAGAGTGTGA 1397
Db      1920 GCAGCAGATGACAGCAGCAGATGAGCAGCAGAGATGACAGCAGAGATGAGCA 1979
Qy      1398 GGAATCTCTGAAATTTATGAGCGAAGAGCTGCGTAGAATGACAGAGATTAATCGATCGT 1457
Db      1980 GCAGCAGATGACAGCAGCAGATGAGCAGCAGAGATGACAGCAGAGATGAGCA 2039
Qy      1458 GAGACAGAGAACTAAGATGACATGAACAGAAACAGGAAAGATGATGACACAGCAG 1517
Db      2040 GCAGCAGATGACAGCAGCAGATGAGCAGCAGAGATGAGCAGAGATGAGCA 2099
Qy      1518 GTTTTATGATTCATCAACAGATCCATGAAGAAAGAGCAGCAAGAGATTT 1577
Db      2100 GAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGCA 2159
Qy      1578 CGAGATTTTGCAGCAGCAGAAAGCTGCCAAGGTTGTTGCGCAGCAGCAGCAGAACTTAA 1637
Db      2160 GAGTGAAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGCAGATGA 2219
Qy      1638 TCCCTTAGCAATGACATGTCGGAAGAGAGCTGAGAGAAAGTCAAGCTTCAATCGATT 1697
Db      2220 GCAGCAGCAGCAGATGAGCAGCAGCAGAGATGAACAGAGAGCAGCAGAGAGCAGGA 2279
Qy      1698 TCAGAAGAAAGAGATGAGAGAGTTTGTGAAGAGAGAGAGATGCTGTAAGAGTCAGA 1757
Db      2280 GCAGCAGAGAGAGAGAGAGAGAGTGAAGAGACAGAGAGAGAGATTAAGAGATCGGA 2339
Qy      1758 GAAGAAATGGAAGACATGAAGAAAGAGCATCAGAGAGATTAATGATCTGAGAAAGA 1817

```

```

Db      2340 GCAGAGATTAGAGGAGCAGAGAGAGATTAGAGAGCAGAGAGAGATTAGAGAGCA 2399
Qy      1818 ATTGATGAGGCTTTGGAACAG 1839
Db      2400 GGAGCAGAGATTAGAGAGAGCAG 2421

```

RESULT 7

```

US-08-770-379-20/C
; Sequence 20, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,379
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-770-379-20

```

Query Match 3.1%; Score 58.8; DB 2; Length 32207;

Best Local Similarity 43.4%; Pred. No. 2.5e-06; Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

```

Qy      1218 GAAGCAGATCTGACATATTCATCACTCTCAAGGCAAAACAGGCTGAAATTCGA 1277
Db      20197 GCAGCAGATGACAGCAGCAGATGAGCAGCAGAGATGAGCAGCAGAGATGAGCA 20138
Qy      1278 GTTGAATCATACCAAGATGTTGTAAAGAGCTGAGGAGATCTTGAGGACATCA 1337
Db      20137 GCAGCAGATGACAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGAGATGAGCA 20078
Qy      1338 GCAGCTGAATCTTTAAGAACAAAGCTCTCAAAACAGAAACAGCAGCCAGAGTGTGA 1397
Db      20077 GCAGCAGATGAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGAGATGAGCA 20018
Qy      1398 GGAATCTCTGAAATTTATGAGGAGAGCTGCGTAGAATGCAAGAGATTAATCGATCGT 1457
Db      20017 GCAGCAGATGACAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGAGATGAGCA 19958

```



```

QY 1458 GAGACAGAGAACTAAGATGACGATGAAACGAAACAGGAAAGATGATGACACAGCAG 1517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19957 GCAGCAGAGATGAGCAGACGAGATGAGCAGCAGCAGATGACAGACAGGATGAGCA 19898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1518 GTTTTCATGATTCATCAATCAACAGATCCATGAAAGAGAGCAGAAAGAGAGATTT 1577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19897 GGAGCAGCAGAGATGAGCAGACGAGATGAGCAGAGCAGACAGATGACAGAGCA 19838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1578 CGAGATGTTGACGACAGAGAAAGCTGCAAGGTTGTTGGCCAGCAGACAGAACTTAA 1637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19837 GATGAGCAGCAGCAGACAGATGACAGCAGCAGATGACAGCAGACAGCAGATGA 19778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1638 TCCCTTACGATGACGATTCGCAAGAGCTGAGAGTCAAGTTCATCGAGTT 1697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19777 GCAGCAGCAGCAGAGATGAGCAGACAGCAGATGAAACGAGACAGAGAGAGAGCA 19718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1698 TCAGAGAGAAAGATGAGAGAGTTTGTGAGAGAGAGAGATCTGTAAAGATCAAGA 1757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19717 GCAGCAGAGAGAGCAGAGAGAGATGAGAGCAGAGCAGAGATGACAGATCAGGA 19658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1758 GAAGAGATGAAAGATGAAAGAGGATCAAGAGAGATATTGATCTGAGAAAGA 1817
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19657 GCAGGAGTTAGAGAGCAGAGAGAGATGAGAGCAGAGCAGAGAGATGAGAGCA 19598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1818 ATTTGATGAGGCTTTGGAACAG 1839
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19597 GGAGCAGAGATGAGAGAGCAG 19576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 8
US-08-757-669A-20/c
; Sequence 20, Application US/08757669A
; Patent No. 6183751

```

GENERAL INFORMATION:

```

; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; NUMBER OF INVENTIONS: 20
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,669A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-757-669A-20

```

```

Query Match      3.1%; Score 58.8; DB 3; Length 32207;
Best Local Similarity 43.4%; Pred. No. 2.5e-06;
Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

```

```

QY 1218 GAAGCAAGATCTGACATATTCATCAATCAACTCTCAAGSCAAAAGCTGTAATTCGA 1277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20197 GCAGCAGATGAGCAGCAGACAGATGAGCAGCAGATGAGCAGCAGATGAGCA 20138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1278 GTTGAATATATCCAAAGATGTTGTAAAGAGCTGAGGCGATCTGAGACATCA 1337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20137 GCAGCAGATGAGCAGCAGACAGATGAGCAGCAGATGAGCAGATGAGCAGATGAGA 20078
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1338 GCAGTGAATCTTCTTAAAGAAAGCTTCAAAACGAACAGACAGCCAGGTCCTGA 1397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20077 GCAGCAGATGAGCAGCAGACAGATGAGCAGCAGATGAGCAGATGAGCAGATGAGA 20018
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1398 GGAATCTCTGAAATTTATGAGCGAGAGAGCTGCTGTAAGCTGAGAGATATTCGATCGT 1457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20017 GCAGCAGATGAGCAGCAGACAGATGAGCAGCAGATGAGCAGATGAGCAGATGAGA 19958
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1458 GAGACAGAGAACTAAGATGACGATGAAACGAAACAGGAAAGATGATGACACAGCAG 1517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19957 GCAGCAGATGAGCAGCAGACAGATGAGCAGCAGATGAGCAGATGAGCAGACAGATGAGA 19898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1518 GTTTTCATGATTCATCAATCAACAGATCCATGAAAGAGACGCAAGAGAGAAATTT 1577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19897 GCAGCAGCAGATGAGCAGACAGACAGATGAGCAGAGATGAGCAGACAGATGAGAGCA 19838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1578 CGAGATGTTGACGACAGCAGAGAAAGCTGCAAGGTTGTTGCCAGCAGCAGACAGAACTTAA 1637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19837 GATGAGCAGCAGCAGACAGATGAGCAGCAGCAGATGAGCAGACAGACAGATGAGCA 19778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1638 TCCCTTACGATGACGATTCGCAAGAGCTGAGAGTCAAGTTCATCGAGTT 1697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19777 GCAGCAGCAGCAGAGATGAGCAGACAGCAGATGAGAAACGAGCAGACAGAGAGAGCA 19718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1698 TCAGAGAGAAAGATGAGAGAGTTTGTGAGAGAGAGAGATCTGTAAAGATCAAGA 1757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19717 GCAGCAGAGAGAGCAGAGACAGAGATGAGAGCAGAGCAGAGATGAGAGATCAGGA 19658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1758 GAAGAGATGAAAGATGAAAGAGGATCAAGAGAGATATTGATCTGAGAAAGA 1817
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19657 GCAGGAGTTAGAGAGCAGAGAGAGATGAGAGCAGAGCAGAGAGATGAGAGAGCA 19598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1818 ATTTGATGAGGCTTTGGAACAG 1839
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19597 GGAGCAGAGATGAGAGAGCAG 19576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 9
US-09-230-371A-20/c
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus

```

US-09-230-371A-20

Query Match 3.1%; Score 58.8; DB 4; Length 32207;
 Best Local Similarity 43.4%; Pred. No. 2.5e-06;
 Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

1218 GAAGCAAGATCTGACATATTCATCACTCTCAAGCAAAACAAAGCTGAATTGCA 1277
 20197 GCAGCAGGATGACACAGCAGAGATGACAGCAGAGATGACAGCAGATGACCA 20138
 1278 GTTGAATCATACCAAGAGTGGTGTAAAGAGCTGAGCAGATCTCTGAGACATCA 1337
 20137 GCAGCAGGATGACAGCAGCAGATGACAGCAGCAGATGACAGCAGCAGATGACCA 20078
 1338 GCAGCTGAACTATTAAAGAAAGCTCTCAAAACAAACAGCAGCCAGGATCTTGA 1397
 20077 GCAGCAGGATGACAGCAGCAGATGACAGCAGCAGATGACAGCAGCAGATGACCA 20018
 1398 GGAATCTCTGAAATTTTGAAGCGAAGCTGCTAAGCTGACAGATTAATCGATCGT 1457
 20017 GCAGCAGGATGACAGCAGCAGATGACAGCAGCAGATGACAGCAGATGACCA 19958
 1458 GAGACAGAACTAAGATGACAGTAAACAGAAACAGGAAAGATGATGACACAGCAG 1517
 19957 GCAGCAGGATGACAGCAGCAGATGACAGCAGCAGATGACAGCAGCAGATGACCA 19898
 1518 GTTTCATGATTCATCAACAGATCCATGAAAGAAAGACGCAAGAGAGAAATT 1577
 19897 GCAGCAGGATGACAGCAGCAGATGACAGCAGATGACAGCAGATGACAGCAGCA 19838
 1578 CGAAGATTTGAGCAGCAGAAAGCTGCCAAGTTTGTGCCAGCAGCAGCAACATTA 1637
 19837 GATATGACAGCAGCAGCAGATGACAGCAGCAGATGACAGCAGCAGATGACCA 19778
 1638 TCCTCTAGCAATGACATGCTCCAAAGAGCTGAGGAATGTCAGATTCAGATT 1697
 19777 GCAGCAGCAGCAGATGACAGCAGCAGCAGATGACAGCAGCAGCAGATGACCA 19718
 1698 TCAAGGAAGAGATGAGAGATTTGTGAAAGAGAGAGATGCTGATAAAGATCAAG 1757
 19717 GCAGCAGGAGCAGCAGCAGATGAGAGAGAGAGATGAGAGAGAGATGAGATCA 19658
 1758 GAAGAGATGAGAGATGAGAGAGATGAGAGATGAGAGATGAGATGAGATGAG 1817
 19657 GCAGGAGTTAGAGAGCAGCAGCAGATGAGAGAGCAGAGAGATGAGAGAGCA 19598
 1818 ATTGATGAGGCTTTGAAACAG 1839
 19597 GCAGCAGGATGAGAGAGCAG 19576

RESULT 10

US-08-781-891-208/c

Sequence 208, Application US/08781891

Patent No. 6090620

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

APPLICANT: Yu, Chang-Bn

APPLICANT: Oshima, Junko

APPLICANT: Mulligan, John T.

APPLICANT: Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

TITLE OF INVENTION: WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,891
 FILING DATE: 27-DEC-1996
 CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: No. 6090620, endburg, Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 240052.419

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ. ID NO: 208:

SEQUENCE CHARACTERISTICS:

LENGTH: 16442 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-781-891-208

Query Match 2.8%; Score 53.2; DB 3; Length 16442;
 Best Local Similarity 47.6%; Pred. No. 7.6e-05;
 Matches 157; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

1472 AGATGCAAGATGAAACAGAGAGAGATGATGATGATGATGATGATGATGATGAT 1531
 16437 AG 16378
 1532 CAATCAACAGATCCATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1591
 16377 AG 16318
 1592 AG 16258
 16317 AG 16108
 1652 ACATTTGCCAAG 1711
 16257 AG 16198
 1712 TGAAGAGATTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1771
 16197 AG 16138
 1772 ACATGAG 1801
 16137 AG 16108

RESULT 11

US-09-618-166-208/c

Sequence 208, Application US/09618166

Patent No. 6583112

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

APPLICANT: Yu, Chang-Bn

APPLICANT: Oshima, Junko

APPLICANT: Mulligan, John T.

APPLICANT: Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

TITLE OF INVENTION: WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group

STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 16442 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 208:
US-09-618-166-208

Query Match 2.8%; Score 53.2; DB 4; Length 16442;
Best Local Similarity 47.6%; Pred. No. 7.6e-05;
Matches 157; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 1472 AGATGAGCATGTAACGAAGAGGAAAGATGATGACACGACGATTTTCATGAT 1531
DB 16437 AGAGAGAGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16378
QY 1532 CAATCAACAGATCCATGAAAGAGAGAGAGAGAGAGAGAGATTTTCAGAT 1591
DB 16377 AG 16318
QY 1592 AGCAG 1651
DB 16317 AG 16258
QY 1652 AGCATTTCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1711
DB 16257 AG 16198
QY 1712 TGGAGAGAGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1771
DB 16197 AG 16138
QY 1772 ACATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATAT 1801
DB 16137 AG 16108

RESULT 12
US-08-781-891-209/c
Sequence 209, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-209

Query Match 2.7%; Score 49.8; DB 3; Length 51259;
Best Local Similarity 50.5%; Pred. No. 0.0015;
Matches 146; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

QY 1525 ATGATTCATCAATCAACAGATCCATGAAAGAGAGAGAGAGAGAGAGAT 1583
DB 290 ATGCTTAAACACACAAATAATACATGAAAGAGAGAGAGAGAGAGAGAGAG 231
QY 1584 GTTGCAG 1643
DB 230 GAG 171
QY 1644 TAGCAATGACGATTGCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1703
DB 170 GCAG 111
QY 1704 GAAAGATGAGAGAGAGATTTGTGAAAGAGAGAGAGAGAGAGAGATGAT 1763
DB 110 GGAGGGGCGAG 51
QY 1764 GATGGAAGCATGAAGAGAGAGAGAGAGAGAGAGAGAGAGATATTTGAT 1812
DB 50 GAG 2

RESULT 13
US-09-618-166-209/c
Sequence 209, Application US/09618166
Patent No. 6583112
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McMaister, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 209:
US-09-618-166-209

Query Match          2.7%; Score 49.8; DB 4; Length 51259;
Best Local Similarity 50.5%; Pred. No. 0.0015; Indels 1; Gaps 1;
Matches 146; Conservative 0; Mismatches 142;

QY 1525 ATGATTCATCAACAGATCATGAAAGA-GACGCAAGAGAGAGATTTCAGAT 1583
DB 290 ATGCTTAAACACAAATATCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 231
QY 1584 GTTGACAGCAGAGAACTGCCAAGTTTGTGGCCACACAGCAGCAATTAATCCCTC 1643
DB 230 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 171
QY 1644 TAGCAATACATTCGCCAAGAGAGCTGAGAAAGTGTCAAGCTTCATCGATTCAAGA 1703
DB 170 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 111
QY 1704 GAAAGAGATGAGAGAGTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1763
DB 110 GAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 51
QY 1764 GATGAGAGCATGAAAGAGAGGATCATGAGAGAGATATTGATCTGAG 1812
DB 50 GGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2

RESULT 14
US-09-177-325-2
Sequence 2, Application US/09177325B
Patent No. 6214983
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sakin, Ugur
APPLICANT: Pfreundschuh, Michael
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins,
FILE REFERENCE: LUD 5525
CURRENT APPLICATION NUMBER: US/09/177,325B
CURRENT FILING DATE: 1998-10-22
NUMBER OF SEQ ID NOS: 15
SEQ ID NO 2
LENGTH: 1276
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
US-09-177-325-2

Query Match          2.5%; Score 47.4; DB 3; Length 1276;
Best Local Similarity 47.2%; Pred. No. 0.00088; Indels 0; Gaps 0;
Matches 144; Conservative 0; Mismatches 161;

QY 1516 AGGTTTTCATGATTCATCAATCAACAGATCCATGAAAGAGAGCGCAAGAGAGAGAT 1575
DB 607 ATGTCTTAAAGAAATATCATGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
QY 1576 TTGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1635
DB 667 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726
QY 1636 AATCCCTTAGCAATGATGATTCGCCAAGAGAGAGAGAGAGAGAGAGAGAG 1695
DB 727 TAATAGTAGTAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786
QY 1696 TTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1755
DB 787 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
QY 1756 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815
DB 847 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906
QY 1816 GAATT 1820
DB 907 AATT 911
```

```

RESULT 15
US-09-411-812A-2
Sequence 2, Application US/09411812A
Patent No. 6261778
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sakin, Ugur
APPLICANT: Pfreundschuh, Michael
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins, And
FILE REFERENCE: LUD 5525.1 CIP
CURRENT APPLICATION NUMBER: US/09/411,812A
CURRENT FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: US 09/177,325
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 2
LENGTH: 1276
TYPE: DNA
ORGANISM: Homo sapiens
US-09-411-812A-2

Query Match          2.5%; Score 47.4; DB 3; Length 1276;
Best Local Similarity 47.2%; Pred. No. 0.00088; Indels 0; Gaps 0;
Matches 144; Conservative 0; Mismatches 161;

QY 1516 AGGTTTTCATGATTCATCAATCAACAGATCCATGAAAGAGAGAGCGCAAGAGAGAGAT 1575
DB 607 ATGTCTTAAAGAAATATCATGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
QY 1576 TTGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1635
DB 667 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726
QY 1636 AATCCCTTAGCAATGATGATTCGCCAAGAGAGAGAGAGAGAGAGAGAGAG 1695
DB 727 TAATAGTAGTAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786
QY 1696 TTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1755
DB 787 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
QY 1756 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815
DB 847 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906
QY 1816 GAATT 1820
DB 907 AATT 911
```

Fri Apr 9 15:36:15 2004

us-10-030-829-2.rml

Page 9

Search completed: April 8, 2004, 16:29:41
job time : 96.7349 secs

This Page Blank (uspto)


```

Db      572 TTTTGAGACTCGAAAAACGAACAGTGTTCAAAGAGTTCTTTGAGATGTAATTCATT 631
Qy      642 GTGATCGACAGATAATTAATACCAAGAGCGAGTGCATTGTTCAGCTTGTCAAGACG 701
Db      632 GACTTTGGAACAAATTAATTAACAACTCGCAATGGCATTTGCCCGCATGCAAAAATG 691
Qy      702 ACCGTGGCCATGATTTGTATATA---CCTGACCCCTCTACTAGTCTATGCGAGACAA 758
Db      692 ACCGTGGCGCATTTGATCTGATCAAGAGGCTGACGCTTTTGTAGTCAATGCTTGAACAA 751
Qy      759 AGGAGCTAGGCGAGTTAAGCTCCATGAGAAATTTGGCTGAAGTTTGAAGAAAGATCTACA 818
Db      752 GGGTTCTACAAAGGTTAACTTCACAGGGAATTTGGCTGATCTCTGGAAGAGAGCTTTC 811
Qy      819 GATGAAAGGCGCATCTGTCAATTCCTTGTGTGAGATTTATGAGCAGTGAAGGTTTGG 878
Db      812 TCGCAGGGGAATCTCAAGTGTACCGCGGGTGAACAAATTTGGAAATGGAAGAGCTG-- 869
Qy      879 TGAGATGAAAAGATTAATTAATTTCTGCTCCCAATGCTCATCATCATGTAATCTAG 938
Db      870 -CAAGAAAGACATGATCGTGAAGTATGTGTGCGCACCAATGTATTGTATGAAATACCTT 928
Qy      939 ACTGATTAAGAGACATTAACGATTAAGTGTCTGCGCATGCGCAACCAAGCTGCTGAATA 998
Db      929 CTGGAAGAAAAGATGAAGATATTAAGTGAAGGCGCATGGAACCAAGAGCTCTTGATTA 988
Qy      999 CTTCGACAAATGAGAGCTCTTGAAGCAGCCCATTTCTATGTGTCCACAGGCGCATGTGTG 1058
Db      989 TTTTGAAGATAATGAAGCAAGTAAGCAACGTATGCTTGTGTGCTCTGTGGGACCGTGTG 1048
Qy      1059 GATGAGTGTCTGATGTTTGAAGAGAGTGCACATGCTGCTATTTTGGAGCGCAACGCTTCA 1118
Db      1049 TATGACCGTGTAAATTTGAAGCTCAGCTGTGGCTATCATGGAAGCTGAAGCTGTGCA 1108
Qy      1119 CCGGAGATTAGCTGAGATGGGTTAAGATGAATGCTGTGGAGTCA--GAAGGCAAGTAT 1175
Db      1109 TAAACATTTGTTAACTCAAGGTACAGAGGAATTCATGCACTTACGCAAGGTTGATTT 1168
Qy      1176 GTTTTCTGAGAGTGTTCGCCAATGTATGTGCTTCTTGAAGACAGACAGATTTGACAT 1235
Db      1169 TGTGCTGTGGGAAAAAGCACTATATATGTTCTTGAACAAAGAGATTAATGAGGC 1228
Qy      1236 ATTCATCAACAATCTCAAGGCAAAACAAGGCTGAATTCAGATTAATCATACCAAGA 1295
Db      1229 ATTTAACAGCAATTTGCCATGTGTAAGCCCTGTGAAGTATGATGATGAGTCTTACATGA 1288
Qy      1296 GATGTTTGAAGAGAGCTGAGAGATCTTCTGAAGCAATCAGCAGCTGAATCTTAA 1355
Db      1289 GATGTGTGTGATTCATATGTAACAGATGAGTGAAGCAACCAACTGAATTAACCTGAA 1348
Qy      1356 GAAACAGCTCTCAAAACAGAACAGAACGCCAAGTGTCTTGAAGATCTCTGAAATTTAT 1415
Db      1349 GAACAGATGTGTTAACAAGAGCAGCAAGTCTTAAAGATGAAGAAATCATTTGGTGTGT 1408
Qy      1416 GAGCGAAGAGCTGCTGAAGTCTGAGAGATTAATCCGATCGTGAAGACAGAACTAAGAT 1475
Db      1409 TACCCAGAGAGCTACGTGAAGCTATTTGAAGAAATATTTTGTGAAGATGAAGCAAAA 1468
Qy      1476 GCAGCATGAACAGACAGAGAGAGATGATGACACAGCAGGTTTTTCATGATTCAT 1535
Db      1469 GAGGACATGAGATGAGAGAGAGATGAAGTCCAGAGAGAAATTTTTCATGTCTAAT 1528
Qy      1536 CAAACAGATCCATGAAAGAGAACGCCAAGAGAGAGAAATTTGAAATGTGAAGAGCA 1595
Db      1529 TGAAGATTAATTCACAAAGCCACAGAGATTAAGAAAGAGGTTTGAAGATTTGAGAGA 1588
Qy      1596 GGAACGTGCAAGAGTGTGTGGCAGAGACAGACAGATTAATCCTCTAGCAATGACGA 1655
Db      1589 AAGAGCCTCAAGGCGCAGAGC-----TTTGAATTTGATTTCTGGAACATTAAGAA 1639
Qy      1656 TTGCCGAAAGAGAGCTGAGAGAGTGTCAAGCTTCATGAGTTTCAAGAGAAAGATGGA 1715

```

```

Db      1640 TCGTAGTAAAGAAATATGTGCAGAGATTCATGATTCACCAAGTCAAGATGTGGC 1699
Qy      1716 GGAAGTTTGTGAAGAGAGGAATGTCTGATAAAGATCAAGGAAGAAAGATGGAAGACT 1775
Db      1700 AGAGTTTGAAGTTGAGCGAGATGATCAAAATCCACGAGGACAGAAAGTTGAAGCT 1759
Qy      1776 GAAGAAAGAGCATCAACGAGAGATATTTGATCTGGAAGAAAGATTTGAAGCTTTTGA 1835
Db      1760 CAAGAAAGAAATACATGATCTGTGAGACTGAGCTGGAAGAAAGAGTTTGAAGCAGCCTTGC 1819
Qy      1836 ACAGCTCATGTACAAGCAT 1854
Db      1820 GGGCTGATGAGAAAGCAT 1838

RESULT 2
US-10-424-599-90522/c
; Sequence 90522, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 90522
; LENGTH: 2218
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2218)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_52754C.1
US-10-424-599-90522

Query Match      21.3%; Score 400.4; DB 12; Length 2218;
Best Local Similarity 67.1%; Pred. No. 1.3e-105;
Matches 615; Conservative 0; Mismatches 291; Indels 10; Gaps 3;

Qy      455 AGGAGTTCTCGACGTGAGATGATGTGATATGCTTCTGAGAGAGAAATGATTCGG 514
Db      2153 AAGAGTCCCTTATGAAGATGATGATGATGAAGAGGAGATGAAGAGAGGTTTG 2094
Qy      515 ATGCTTTGGATGAT-TCTGATGAGACCTTGCAAGTGAATGATGATCGGATGAGT 573
Db      2093 ATGCTATGAGATACCGAGATGATATCTAATGATGATGATTAATGATTCGATATCTAGT 2034
Qy      574 CAAAGAGCAGATGATCAAGAAAGAGAAATAGTGTCAAAAGTCTTTGGCAGCTTG 633
Db      2033 CAAAGAAACAGAACTCGTAAGAAAGCAAAATGTTTAAGATTTCTTTGAAGAACTTG 1974
Qy      634 GATAGCTGTGATGAGCAGATTAATTAACCAAGAGCGAGTGCATTTGTCAGCTTGT 693
Db      1973 GATGCTTATCTATCGAAAGATCAATGAACCGAAAGGAGTGCATGTAACCTTGT 1914
Qy      694 CAGAAAGGACCTGTGCGCATGATTTGTGATA---ACCTGCAACCTCTACTAGCTGATGCG 750
Db      1913 CGAGGTGTCCCGGTGTATTTGATTTGTACAGAGAGCTGACGCTCTGATTAATCTATGCC 1854
Qy      751 AGGACAAAGAGAGCTAAGCGAGTTAAGCTCATAGAGAAATTTGCTGAAGTTTGAAGAA 810
Db      1853 AAAACAAAGGGGTCAAAAAGGTGAAGATCATAGGAGATTTGCTGACACTTTTGAAGAG 1794
Qy      811 GATCTACAGATGAGAGCGCATCTGTCAATCTCTGTGTGAGATTTATGGCGCATGGAAG 870
Db      1793 GAACTGCCAGAGAGGGGCTGTGCAAGTAATTCACCTGGGGAAGTATTGTGAATGGAAA 1734

```



```

QY 871 GGTGGGATGAGATGAAAGATTAATGATGCTGCGCTCCATGGTCATCATCAG 930
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1733 GGTTTAAAGATGAGAGAAAGATCATGAATATGTTGGCTCCATGGTTCATTCAG 1674
QY 931 AATACTAGACTGATTAAGAGACATTAAGTGGCTCGGCAATGGCAACCAAGACTG 990
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1673 AATACAAAGGCTTGAAACAGGATGAAATGATTAAGTGGCTAGTATGGTAAACGAGACTT 1614
QY 991 CTGGAATCTTGCACAAAGTATGAGGCTCTTAAGACAGCCATTCCTATGCTCCACAGGGC 1050
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1613 CTTAATTAATTTAAGCCATATGATGCTGTAAGAGCTGACACGCTTATGCCCCGAGGT 1554
QY 1051 CATCGGGATGATGATGCTGATGTTTGAAGAGAGAGCCATGGCTATTTGAGGCCGAA 1110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1553 CATCGGGGATGAGCTTTGATATTTGAAGCATCAGCTATAGGTTATCTTGAAGGCTGAG 1494
QY 1111 CGCCTTCCACCGGAGTTAGCTGAGATGGGGTTAGATTAAGTTCCTGGGGTCAAGAGCC 1170
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1493 CGTCTACACAGCACTTTGACAGAAACAGAACTGACCGAGATGCTTGGTTAGTCAACAC 1434
QY 1171 AGTATGTTTT-----CTGAGAGTGTTCGCCCATCTGATAGGCTTCTTGCAAGAGCAA 1224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1433 CGTAGATTAATTTCTCCCTGGTGGCCGCGCACAGCTTATGATTAATGCAATTAAGAA 1374
QY 1225 GATCTGACATATTCATCAACACCTCTCAAGGCAAAAGGCTGAATTCGAGTTGAAA 1284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1373 GACTTGGACTTTTTCACAGACATTCGCCAAGTAAATCTTAGACTCAAAATATGACATGAG 1314
QY 1285 TCATACCAAGAGATGTTGTAAGAGAGCTGAGGAGATCTCTGAGAGCAATCAGACGCTG 1344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1313 TCATACCAAGAGATGTTGTAAGCAACTTCGCAAGAGATGAGGCAACGAGCTGCTT 1254
QY 1345 AACTACTTTTAAGACA 1360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1253 AATTATTTGAAGGAAA 1238

```

RESULT 3

```

US-10-424-599-91177
; Sequence 91177, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yinhua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 91177
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53343C.1
US-10-424-599-91177

```

```

Query Match 13.6%; Score 254.6; DB 12; Length 1377;
Best Local Similarity 66.3%; Pred. No. 3.5e-63;
Matches 417; Conservative 0; Mismatches 199; Indels 13; Gaps 3;
QY 352 TATGATTAACAACCTTTGTGACACCCCACTGTATCTCGCCCTCTTGTGAAGAGAGATGG 411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 29 TATGAAGAACAATATGTGACTCCAAATCTGTAAATCGGCTCCACTCGAGATGGCTGG 88
QY 412 AATGGCAGCAAGAGAGAGGTTTGTCTCAGACACACAGCTGTGACAGAGTTTCTGACGTG 471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 89 AACTGGCAATCCAGACCTGTGTGCAATTCAGTCCAAATGTAAAGGATGAATAATTCACAGAG 148

```

```

QY 472 G-----AGATGATGTGATTAATGCTTCTGAGAGAAAGAAATGATTCGATCTTTG 522
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 149 GACCTTCAAAAAGAAATTAATGATGATGATGATGAGGAGAGAGAAATCTGATGATTTG 208
QY 523 GATGATTTGATGACGACCTTGCAGAGTGAATGATTAATGATCTCGATGTGATGATCAAAAAGAC 582
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 209 GAAGATTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 268
QY 583 CATGATGACGAGAAAGAGAAATTAAGTGTCAAAAAGTCTTTGGCAGCTTGTGATGATG 642
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 269 CATGAGACAGCTTAAGAAAGAAAGAAATGATTAAGAAATTTTGAAGATTTGATGATGCTTG 328
QY 643 TCGATGACAGATTAATTAACACAGAGCAGTGGCATTTGTCAGCTTGTGACAAACGGA 702
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 329 ACTGTTGAACAGATTAACAGAACACAGAGACAGTGGCATTTGTCAGCTTGTGACAAACGGA 388
QY 703 CCGTGTGCTATGATTTGATTAAC--CTGACACCTCTACTAGCTCATGCTGAGAGACAAA 759
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 389 CCGAGGCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 448
QY 760 GGAGCTAGGCGAGTTAAGCTCCATAGAGAAATGAGTGAAGTTTGAAGAAAGATCTACAG 819
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 449 GATCAAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
QY 820 ATGAGAGCGCATCTGTCTCTTGTGTGATTAATGAGCAAGTGAAGGTTTGGGT 879
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 509 AAAAGGGTACTTCAATTAATTCAGCTGGGAGATTTGAAAAGTGAAGGTTTGAAA 568
QY 880 GAGATGAAGAAGATTAATTAATTTGTCTGCTCCATGCTCATCATCAATTAATTAAG 939
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 569 GAAGAAG-AAAGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
QY 940 CTGGAATGAAGAGATTAAGATTAAGTGGCT 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 628 CTGAACAGATGAATAATGATTAAGCGAT 656

```

RESULT 4

```

US-10-424-599-91176
; Sequence 91176, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yinhua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 91176
; LENGTH: 1589
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53342C.1
US-10-424-599-91176

```

```

Query Match 11.5%; Score 215.6; DB 12; Length 1589;
Best Local Similarity 64.6%; Pred. No. 9.1e-52;
Matches 358; Conservative 0; Mismatches 184; Indels 12; Gaps 2;
QY 352 TATGATTAACAACCTTTGTGACACCCCACTGTATCTCGCCCTCTTGTGAAGAGAGATGG 411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 29 TATGAAGAACAATATGTGACTCCAAATCTGTAAATCGGCTCCACTCGAGATGGCTGG 88
QY 412 AATGGCAGCAAGAGAGAGGTTTGTCTCAGACACACAGCTGTGACAGAGTTTCTGACGTG 471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 89 AACTGGCAATCCAGACCTGTGTGCAATTCAGTCCAAATGTAAAGGATGAATAATTCACAGAG 148
QY 472 G-----AGATGATGTGATTAATGCTTCTGAGAGAAAGAAATGATTCGATGCTTTG 522

```

Db 149 GACCTTCAAAAGAAATTATGTGTGTTGATGATGATGGGAGGAGGAGAACTTGATGATTTTG 208

Qy 523 GATGATTTCTGATGAGACGCTTGCAGTGTATGATTAATGACTCGGATGTAGTCACAAAAGC 582

Db 209 GAAGTACTGATGATGATGATCTTGATGATGATACGATCATATTTCTATGTGATGTCAAAAGAGT 268

Qy 583 CATGATTCACGAAGACAGATTAAGTGTTCAAAAAGTTCTTTGGACCTTGGATTAACCTTG 642

Db 269 CATGAGACACGTAAAGAAAAGCAAAATGATTAAGAAAGTCTTTTGAGATTTTGGATGGCTTG 328

Qy 643 TCGATCGACGAGTAAATTAATGAACCAAGAGCGGAGTGTTCAGCTTGTCAAGACGA 702

Db 329 ACTGTTGAACAGATTAACCAACACGAAAGACAGTGGCATTTCCAGCTTGCACAGGTGGT 388

Qy 703 CCTGGTGGCATCGATGTGATATA---ACCTGCACCTCTTACTAGTCAATCGAGAGACAAA 759

Db 389 CCTGGAGCTATTGATTTGATGTCACAGAGACCTGCAGCCCTTGTGACACATGCCAAAACAAAG 448

Qy 760 GGAGCTAAGCGGAGTTAAGCTCCATAGAGAAATTGGTGAAGCTTTTAAGAAAAGGATCTACAG 819

Db 449 GGATCAAAAAGGGTGAAGATTCACAGGGGAGCTTGGTATATCTTTGGATGAGGAATTCACG 508

Qy 820 ATGAGAGGCCCATCTGTCAATCTCTTGTGGTGAAGATTATAGGCACTGTGAGAGGGTTTGAGT 879

Db 509 AGAAGGGGGCTCTGAGTAATTTCCACTCGCGAAGATTTGGTAAGTGAACGGTTTAAAA 568

Qy 880 GAGAGTAAAAAGGA 893

Db 569 GTGCAGCAGAAAAG 582

```

RESULT 5
US-10-424-599-73594/C
Sequence 73594, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223) B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 73594
LENGTH: 2890
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_37470C.1
US-10-424-599-73594

```

Query Match	Best Local Match	Similarity	Score	DB	Length
Matches	284	Conservative	10.58	197.8	2890
			69.48	Pred. No. 28-46	
			0	Mismatches 122	Indels 3; Gaps 1
Qy	548	GTGATGATTATGACTCGATGTGATGTAAGTCMAAAGCCATGATATCAAGAAAGCATTAAGT	607		
Db	2890	GTGATGATTATGATTCGTGATCTAGTCMAATTGACCAAGAGACTCGTAAGAAAGCAAG	2833		
Qy	608	GTTTCAAAAAGTTCTTGGGAGCTTGGATAGCTTGTGCATCGACAGATPAATGAACAC	667		
Db	2830	GGTTTAAGATTCTTTGAGAACCTTGATGACCTATATCAAAAAGATCAATGAACCG	2777		
Qy	668	AGAGGCAAGTGCATTGTCCAGCTTGTACAGAACGACCTGGTGCATCGATTGGTATA---	724		
Db	2770	AAAGCAGTGGCACTGTACAGCTTGTGAGGTGGTCCCGTCTATTGATTGGTTACAG	2711		
Qy	725	ACCTGCACCTCTACTAGCTCATGCGAGACMAAAGAGCTTAAGCCAGTTAACTCCATA	784		

Db 2710 GACTGACGCTCTGTATTACTCATGCCAAAACAAAGGGCTCMAAAAAGGTGAAGATCCATA 2651

Qy 785 GAGAAATGGCTGAATGTTTATGAAAAGANTCTACAGATGAGAGCGCATCTGTCAATTCCTT 844

Db 2650 GCGAGTTTCTGACGTTTGGAAAGAGAACCTGGCGCAAAAAGGGCTCTGCGATTAATTCAC 2591

Qy 845 GTGCTAGATTTATGGCAGATGGAAGGGTTTGGGTGAGATGAAAAGATTTAGAAATTG 904

Db 2590 CTGGGAAATTAATTTGTAAAGTGAAGAGTTTAAAAAGATGATGAAAGATCATGAATAATG 2531

Qy 905 TCTGGGCTCCAAATGTCATCATCATGAATTAATAGATCTGATTAAGGACGA 953

Db 2530 TTTCGGCTTCGATGTTGTTCATTCACATACAAAGGCTTGAACAGCCTGA 2482

```

RESULT 6
US-10-424-599-36215/c
; Sequence 36215, Application US/10424599
; Publication No. US20040031072A1
GENERAL INFORMATION:
; APPLICANT: Ia Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 36215
; LENGTH: 1103
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1103)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_132704C.1
US-10-424-599-36215

```

Query Match	Best Local Similarity	9.3%;	Score 175.4;	DB 12;	Length 1103;
Matches 285;	Conservative	65.7%;	Pred. No. 3.9e-40;	Mismatches 147;	Indels 2;
					Gaps 2;
QY	1172	GTATGTTTCTGAGAGTGTGCGCAACTGTATGAGCTTCTCTTGCAACGACGAGATCTGG	1231		
Db	434	GTGGGCTCCCTGGGGGGAATGGCCAGCTCTATGTATACATGGCTTACAAAAGAAATATGG	375		
QY	1232	ACATATTCAATCAACACTCTCAAGGCAAAAAGGCTGGAATTGAGTGAATCAATACC	1291		
Db	374	ACTTTTTCACACACACTTCCCAAGGTAAATCTAGACTCAAAATATGACATAAGATCATATC	315		
QY	1292	AAGAG-ATGGTGTGAAGAAGCTGAGGCAAGTCTTGAGACAAATCAGAGCTGAATAC	1350		
Db	314	AGAGGAATGTTTTCACCCAAATTCGGCAAAATGAATGAAGCAACCAACACTGCTTAT	255		
QY	1351	TTTAAAGAACAGCTCTCCAAAACAGAAACAGACGCGCAAGGCTTGAAGATCTCTGGAA	1410		
Db	254	TTTAAAGAACAGGCTGTGAAAGGTGCACAAACACAAAAGCTTTGAAGAAATTAATGGT	195		
QY	1411	ATTATGAGCAGAGAAGCTGCGTGAATCTGCAGAGATTAATCGATCTGTGACAG-AGAAC	1469		
Db	194	ATAATGCTGAGAGAGCTGGCGGAAGCAATGAGAGAAATTCGATTGTGAGGAGGAGAAC	135		
QY	1470	TAAAGTCAGATCAACAGAACAGGGAAGAAGATGATGCACAGCAGGTTTTCATGGA	1528		
Db	134	TAAATGCAATCATGACAGACCAAAAAGNAGATGTTATGCAAGAACAAITCTTCAAGGA	75		
QY	1530	TTTCATCAACAGATCCATGATAAAGAGAGCGCAAGAGAGAGAAATTCAGATGTTGCA	158		
Db	74	CCAGATCAAAATCTTATGATATTCACAGGATGCAAGAGAAAGAAATTCGAGAGAGATGC	15		

Qy	1590	GCAGCAGGAACGTG	1603
Db	14	ACAGGAGGAACGTG	1

RESULT 7
ITS-10-A2

```

US-10-424-599-73591/c
: Sequence 73591, Application US/10424599
: Publication No. US20040031072A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 73591
: LENGTH: 1375
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(1375)
: OTHER INFORMATION: unsure at all n locations
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT3847_37468C.1
US-10-424-599-73591

```

Query Match	9.0%;	Score 169.8;	DB 12;	Length 1375;
Best Local Similarity	63.0%;	Pred. No. 1.9e-38;		
Matches 247; Conservative	0;	Mismatches 108;	Indels 3;	Gaps 1;

QY	548	GTATNGATTATGACCTCGGATGAGTCAAAAGAGCATGANTCAGAAAGCAGATTAAGT	607
Db	1375	GTGATGATTAATGATTCGTATATCTAGTCAAAAGAGCCAGAGACTCTGTAAAGAGACAAT	1316
QY	608	GGTTCAAAAGATTCCTTTGGCAGCTTGAGATAGCTTTCGATCGAGCAGATTAATGAAACCA	667
Db	1315	GGTTTAAAGATTTCTTTAGAACTTTGATGGCTATCTATATGAAAAGATCAATGAACGG	1255
QY	668	AGAGCAGATGTCATTGTCACGCTGTCAAAACGAACTGTGTCATCGATTGGTATA---	724
Db	1255	AAAGGCAATGCGCACCGCTCAGCTTTCGAGGAGTGTCGGGTGTCATTTGATTTGATACAG	1199
QY	725	ACCTGCACCCCTCTACTAGCTCATGCGAGACAAAAGAGCTAAGCGAATTAACTCCATA	784
Db	1195	GACTGCAGACCTCTGATTTACTCATGCCAAAACAAAGGGGTCAAAAAGGGTGAAGATCCATA	1133
QY	785	GAGAAATTGGCTGAAGTTTAAAGAAAGATCTACAGATGAGAGCGGCANTCTGTCAATTCCT	844
Db	1135	GGGAGTTTGTCTGAGCTTTTGGAGAGGAATCTGCCGAAAAGGGGCTCTCAGATTAATTCAC	1076
QY	845	GTGATGAGATTATTAAGCGAGCTGGAAGGGTTTGGGTGAGATGAAAAGATTAATGAAT	902
Db	1075	CTGAGGAAGTATTTGTGTAAGTGAAAGGTTTAAAGATGAGGGGAAAAGATCTATGGAAT	1018

RESULT 8
ITE-10-42

US-10-424-599-99770/c
Sequence 99770, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```

; FILE REFERENCE: 38-21(533223) B
; CURRENT APPLICATION NUMBER: 05/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 99770
; LENGTH: 861

```

; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_61108C.1
 US-10-424-599-99770

Query Match	Best Local Similarity	5.8%;	Score 109.6;	DB 12;	Length 951;
Matches	313;	Conservative	0;	Mismatches 219;	Indels 42;
					Gaps 3;
Qy	1280	TGAATCATATCAAGAGATGGTGTAAAGAGCTGAGGCGAGATCTTGGAGACATACGC	1338		
Db	949	TGAATCATATCAAGAGATGGTGTAGAGAAAGATATTAATATGATATGTCCGA	890		
Qy	1340	AGCTGAATCTCTTAAAGACAAGCTCTCAAAACGAACAAGCAGCCAGGTGCTGAG	1398		
Db	889	AGCTTGAATATTAAGAGCATGTGTAGCCAAAGAACAAATCAATCACAAATGCTGAG	830		
Qy	1400	AATCTCTGAATATTAATGAGCAGAGAGCTGCGTAGAATGTCAGAGATTAATCGATGA	1458		
Db	829	ACTCTTTATATTAATCAAGTAGAAGATTAAGCATGACAACTGAAAAAACCTGTGTGACC	770		
Qy	1460	GACGAGACCTAAGATGACGATGACATGAACGAACAGGAAAGATGATGACACACAGCT	1518		
Db	769	AACGAGACA-----AGGATATGATGACATTAAGAAAAA	737		
Qy	1520	TTTTCATGATTCATCAACACAGATCCATGAAAGAGAGCGCAAGAGAGATTCG	1578		
Db	736	ATTTCCAGAGCCAAATCCAGACACTTTCAACAGCCATAGCTGCAAAGAGATTAAGTTG	677		
Qy	1580	AGATGTTGCAGCAGCAGGAACTGCCAAGTTGTTGCCACAGCAGCAGCAATTAATC	1638		
Db	676	TGAATTTACAGCTGCAG-----GCAATGCAAGAAAAAGTAAAGAGCTCTGCAGAAAGAT	623		
Qy	1640	CCTCTAGCAATGACGATTTGCCGAAGAGAGCTGAGAAAGTCAAGCTCATTCAGATTC	1698		
Db	622	CTTCTGAGAAAGAGAC-----AAGGTGAGAAATATCTTATGATTTCTGAAAGCTC	572		
Qy	1700	AAGAGAAAGATGAGAGAGTTTGTGAAGAAGAGGAGATCTGATTAAGAGTCAAGAGA	1758		
Db	571	AAGCAAGAGATGAACCAATTCAGAGCAGAGAGAGAGAAATTTAAAAAAATTCATGAG	512		
Qy	1760	AGAAAGATGAAGACATGAAGAGAGCATCAAGAGAGATATTTGATCTGAGAAAGAT	1818		
Db	511	AGAAAAAGTTGGCTCTGAAGAAAGAGAGTGGCAAGAGACAGTGTGAACCTTGAGAAAGAGT	452		
Qy	1820	TTGATGAGGCTTTGGAACAGCTCATATGTAACAAGCA	1853		
Db	451	TAGGAATGAATCAGCGAGCTCATGGAACAAGTA	418		

RESULT 9
116-09-29

```

US-09-294-093B-1296
Sequence 1296, Application US/09294093B
Patent No. US20010051335A1
GENERAL INFORMATION:
APPLICANT: Lalagudi, Ragunath, V.
APPLICANT: Ito, Laura, Y.
APPLICANT: Sherman, Bradley, K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSSEL
FILE REFERENCE: PL-0009 US
CURRENT APPLICATION NUMBER: US/09/294,093B
CURRENT FILING DATE: 1999-04-16
PRIORITY APPLICATION NUMBER: 60/082,567
PRIORITY FILING DATE: April 21, 1998
NUMBER OF SEQ ID NOS: 6207
SOFTWARE: PERL Program

```


Sequence 208, Application US/10374077
Publication No. US20040006779A1
GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui
Oshima, Junko

Mulligan, John T.
Schellenberg, Gerald D.

TITLE OF INVENTION: ANTIBODIES AGAINST GENE PRODUCTS RELATED TO
WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group

STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,077

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Roseman, Stephen

REGISTRATION NUMBER: 43,058

REFERENCE/DOCKET NUMBER: 100107.401D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 682-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 208:

SEQUENCE CHARACTERISTICS:

LENGTH: 16442 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 208:

US-10-374-077-208

Query Match

Best Local Similarity 47.6%; Score 53.2; DB 15; Length 16442;

Matches 157; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

1472 AGATGACGATGAAACGAGCAGGGAAGATGATGACACGACGATTTTTCATGATT 1531

16437 AGGAGGAGGAGCGAGCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16378

1532 CAATCAACATGATTCATGAAAGAGAGCAGCAAGAGGAGGAGGAGGAGGAGGAGG 1591

16377 AGCAGGAGCGAGCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16318

1592 AGCAGGAGCGTGCAGAGTTTGTGGCCACGACGAGGAGGAGGAGGAGGAGGAGGAG 1651

16317 AGGAGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16258

1652 ACCATTTCCGAAAGAGAGTGAAGAGTGTCAAGTTTCATGAGTTTCAAGAGAGAG 1711

16257 AGGAGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16198

1712 TGGAGGAGTTTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1771

16197 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16138

1772 ACATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1801

16137 AGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16108

RESULT 15

US-10-108-260A-1966

Sequence 1966, Application US/10108260A

Publication No. US20040005560A1

GENERAL INFORMATION:

APPLICANT: HELIX RESEARCH INSTITUTE

TITLE OF INVENTION: No. US20040005560A1 full length cDNA

FILE REFERENCE: H1-A0106

CURRENT FILING DATE: 2002-03-27

NUMBER OF SEQ ID NOS: 5458

SOFTWARE: Patent Ver. 2.1

SEQ ID NO 1966

LENGTH: 2479

TYPE: DNA

ORGANISM: Homo sapiens

US-10-108-260A-1966

Query Match

Best Local Similarity 43.9%; Score 52.8; DB 15; Length 2479;

Matches 225; Conservative 0; Mismatches 287; Indels 0; Gaps 0;

1328 AGGACATTCGACGCTGAACTTAAGAACCAAGCTCTCAAAACAGAACAGACGCCA 1387

1377 AGGAGAAAGATGTGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1436

1388 AGTGTCTGAGGAATCTCTGAAATTATGAGCGAGAACTGCTGAACTGAGAGGATA 1447

1437 AGGAGAAAGCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1496

1448 ATCGATCTGAGACAGAGAACTTAAGTGCAGATGAAACAGAACAGGAGGAGGAGGAG 1507

1497 AGGAGATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1556

1508 CACACGACAGGTTTTCATGATTCATCAACAGATCCATGAAAGAGAGAGGAGGAGGAG 1567

1557 AGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1616

1568 AGGAGAAATTCGAGATGTTGTCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1627

1617 AGGAGAAATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1676

1628 AGAATTAATCCCTCTAGCAATGACGATTCGCAAGAGAGGAGGAGGAGGAGGAGGAG 1687

1677 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1736

1688 TCATCGAGTTTCAAGAGAAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1747

1737 AGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1796

1748 AAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1807

1797 GGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1856

1808 TGGAGAAAGATTTTGAAGGCTTTTGAACAG 1839

1857 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1888

Search completed: April 8, 2004, 20:41:55
Job time : 471.907 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 19:27:05 ; Search time 60 Seconds
(without alignments)
2943.203 Million cell updates/sec

Title: US-10-030-829-3

Perfect score: 3313

Sequence: 1 MSSRAGPMSEKKVQGGYRP.....EPDEALEQMLYKHLNEDD 625

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database: A_Geneseq_29Jan04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3313	100.0	625	4 AAB31798	Aab31798 Amino ac
2	191	5.8	652	6 ABP98826	Abp98826 Human str
3	182	5.5	1972	2 AAM81171	Aam81171 Human BAZ
4	182	5.5	1972	6 ABR64241	Abt64241 Angiogene
5	179.5	5.4	615	6 AAE32117	Aae32117 Human cyt
6	179.5	5.4	1372	2 AAM56473	Aaw56473 Protein w
7	179	5.4	1203	4 AAM79264	Aam79264 Human pro
8	178	5.4	1203	5 ABG96296	Abg96296 Human ova
9	175	5.3	1879	4 AAM25750	Aam25750 Human pro
10	174.5	5.3	580	6 ABR57420	Abt57420 Human NOV
11	172.5	5.2	593	5 ABP62858	Abp62858 Human pol
12	172.5	5.2	933	4 ABG08600	Abg08600 Novel hum
13	172	5.2	892	7 ADB79862	Adb79862 Rat myosi
14	172	5.2	892	7 ADD47859	Add47859 Rat Prote
15	172	5.2	892	7 ADD47855	Add47855 Rat Prote
16	171.5	5.2	593	6 ABR57419	Abt57419 Human NOV
17	170.5	5.1	676	7 ADA08247	Ada08247 Human tum
18	170.5	5.1	725	2 AAM39165	Aaw39165 Human RHA
19	170.5	5.1	725	5 ABG60842	Abg60842 Human rec
20	170.5	5.1	725	5 AAU11436	Aau11436 Human hya
21	170.5	5.1	725	6 ABR57021	Abt57021 Human RHA
22	170.5	5.1	725	7 ADA08245	Ada08245 Human tum
23	170.5	5.1	725	7 ADC02450	Adc02450 Human rec
24	168.5	5.1	1192	3 AAB18165	Aab18165 Plasmodin
25	167.5	5.1	1238	4 ABB62022	Abb62022 Drosophi

26	167.5	5.1	1294	4 ABB63502	Abb63502 Drosophi
27	167.5	5.1	1898	2 AAy30795	Aay30795 A human t
28	167.5	5.1	1898	7 ADD48869	Add48869 Human Pro
29	167.5	5.1	2274	4 ABB58657	Abb58657 Drosophi
30	167.5	5.1	3899	6 ABR32048	Abt32048 Human cer
31	167.5	5.1	3917	6 ABR92050	Abt92050 Human cer
32	166.5	5.0	1979	3 AAB18171	Aab18171 Plasmodin
33	166	5.0	949	4 ABG15508	Abg15508 Novel hum
34	166	5.0	1162	3 AAY96255	Aay96255 Kapost's
35	166	5.0	1162	3 AAY58500	Aay58500 HIV8 ORF
36	166	5.0	1162	4 AAB62331	Aab62331 Amino ac
37	166	5.0	1162	5 ABB05621	Abb05621 Kapost's
38	165	5.0	661	6 ABP98842	Abp98842 Human str
39	165	5.0	2633	4 ABG06505	Abg06505 Novel hum
40	165	5.0	2663	4 AAM39097	Aam39097 Human pol
41	165	5.0	2668	4 AAM40883	Aam40883 Human pol
42	164	5.0	1354	2 AAM23654	Aaw23654 Physiol
43	164	5.0	1354	2 AAM71020	Aaw71020 A modifie
44	164	5.0	1354	2 AAY07082	Aay07082 Renal can
45	164	5.0	1354	6 ABU05182	Abu05182 Human exp

ALIGNMENTS

RESULT 1
ID AAB31798 standard; protein; 625 AA.
XX
AC AAB31798;
XX
DT 15-MAY-2001 (first entry)
XX
DE Amino acid sequence of the Arabidopsis SGS3 polypeptide.
XX
KM SGS3 gene; post-transcriptional inactivation; RNA degradation;
XX viral resistance; resistance; fatty acid content; protein content.
XX
OS Arabidopsis thaliana.
XX
PN WO200105951-A2.
XX
PD 25-JAN-2001.
XX
PF 13-JUL-2000; 2000MO-FR002052.
XX
PR 16-JUL-1999; 99FR-00009417.
XX
PR 26-JAN-2000; 2000FR-00001006.
XX
PA (AVET) AVENTIS CROPS SCIENCE SA.
XX (INRG) INST NAT RECH AGRONOMIQUE.
XX
PI Beclin C, Elmayan T, Vaucheret H;
XX
DR WPI; 2001-159529/16.
XX N-PSDB; AAF25374.
XX
PT New SGS3 gene from Arabidopsis thaliana, useful for increasing virus
XX resistance in plants and, when inhibited, for increasing transgene
XX expression.
XX
PS Claim 8; Fig 1; 36pp; French.
XX
CC The present sequence represents an Arabidopsis thaliana SGS3 polypeptide.
XX The SGS3 gene is essential for post-transcriptional inactivation
XX (degradation of RNA) and for resistance to viruses. Overexpression of
XX SGS3 results in plants with increased resistance to viruses, while
XX inactivation of SGS3 in transgenic plants (e.g. by expressing antisense
XX RNA, by mutation or by homologous recombination) increases the level of
XX the transgene product. This product may e.g. impart resistance (to
XX herbicide, insects or pathogens), alter contents of essential fatty acids
XX or proteins, or is pharmaceutically active, e.g. an immunoglobulin or
XX interferon

XX Sequence 625 AA;
 SQ Query Match 100.0%; Score 3313; DB 4; Length 625;
 Best Local Similarity 100.0%; Pred. No. 1.9e-250; Indels 0; Gaps 0;
 Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRAGPMSEKKEKVVQGGYRPEVQLVQGLAGTTLASSQDDGGEMVEVSKKKKKPGNTSG 60
 DB 1 MSSRAGPMSEKKEKVVQGGYRPEVQLVQGLAGTTLASSQDDGGEMVEVSKKKKKPGNTSG 60
 QY 61 KTWVSONSNPPRAGWGQOQGRGNNVSGRGNNGRGIOANISGRGRLSKRYDN 120
 DB 61 KTWVSONSNPPRAGWGQOQGRGNNVSGRGNNGRGIOANISGRGRLSKRYDN 120
 QY 121 NFVAPPVSRPPLLEGKNNMARGSAQHTVQEPPEVDVDVNASSEENDSDALDSDDD 180
 DB 121 NFVAPPVSRPPLLEGKNNMARGSAQHTVQEPPEVDVDVNASSEENDSDALDSDDD 180
 QY 181 LASDDVDVQSHSGSRKQNKWPKFPGLSLSLSEIQLNEPORQWHCPACQNGPAIDW 240
 DB 181 LASDDVDVQSHSGSRKQNKWPKFPGLSLSLSEIQLNEPORQWHCPACQNGPAIDW 240
 QY 241 YNHPLLAHARTGARRVKLHRELAIVLEKOLQMGASVIPCSEITYQWKGLGEDEKDYK 300
 DB 241 YNHPLLAHARTGARRVKLHRELAIVLEKOLQMGASVIPCSEITYQWKGLGEDEKDYK 300
 QY 301 IYWPVWYIWNTRLDKDDNDKMLGMGNQELLEYFDKYEARARHSYGGPGRGMSVLMFE 360
 DB 301 IYWPVWYIWNTRLDKDDNDKMLGMGNQELLEYFDKYEARARHSYGGPGRGMSVLMFE 360
 QY 361 SSATGYLAEARLHRELAEMGLDRIANGQKSMFSGVRLYGLATKODLIDFNQSHQK 420
 DB 361 SSATGYLAEARLHRELAEMGLDRIANGQKSMFSGVRLYGLATKODLIDFNQSHQK 420
 QY 421 TRKFEKLSQEVNKKELROISDNOOLNFKKLSKQNHAVYLESLFISEKIRRTA 480
 DB 421 TRKFEKLSQEVNKKELROISDNOOLNFKKLSKQNHAVYLESLFISEKIRRTA 480
 QY 481 EDRIRYORTKMOHEONREEMDAHDPFMDSIKQIHERRPAKENEEMQOQORAVVQ 540
 DB 481 EDRIRYORTKMOHEONREEMDAHDPFMDSIKQIHERRPAKENEEMQOQORAVVQ 540
 QY 541 QOQINPSSNDCKRAEEVSSFLFEQEKMEEFVEREMLIKQEKMEEMKKRHEE 600
 DB 541 QOQINPSSNDCKRAEEVSSFLFEQEKMEEFVEREMLIKQEKMEEMKKRHEE 600
 QY 601 FDLKEPFDEALBQMTKGLHNEED 625
 DB 601 FDLKEPFDEALBQMTKGLHNEED 625

RESULT 2
 ID ABP98826 standard; protein; 652 AA.
 AC ABP98826;
 DT 15-JUL-2003 (first entry)
 DE Human structural and cytoskeletal associated protein #17.
 XX Cytosolic; antiarteriosclerotic; anticonvulsant; nootropic; antianginal;
 XX neuroprotective; cerebroprotective; hypotensive; cardiact; osteopathic;
 XX antiinflammatory; antiarthritic; virocid; gene therapy; human; stroke;
 XX structural and cytoskeleton-associated protein; SCAP; cancer; angina;
 XX heart failure; osteoporosis; Huntington's disease; hypertension;
 XX Homo sapiens.
 XX MO2003031940-A2.
 XX

PD 17-APR-2003.
 XX 10-OCT-2002; 2002MO-US032851.
 PF 12-OCT-2001; 2001US-0328931P.
 PR 19-OCT-2001; 2001US-0360681P.
 PR 02-NOV-2001; 2001US-0343896P.
 PR 09-NOV-2001; 2001US-0346308P.
 PR 16-NOV-2001; 2001US-0332385P.
 PR 07-DEC-2001; 2001US-0340776P.
 PR 11-JAN-2002; 2002US-0347703P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Becha SD, Bhattacharya U, Borowsky ML, Burrill JD, Chang H,
 PI Chawla NK, Elliott VS, Emerling BM, Forsythe IJ, Gorrad AE, Lal PG,
 PI Griffin JA, Halalia AJ, Ho A, Ison CH, Kable AE, Khatre R, Luo W,
 PI Lee S, Lee EA, Lee SY, Lehr-Mason PM, Li JX, Lindquist EA, Luo W,
 PI Marquis JP, Ramkumar J, Richardson TW, Sprague MW, Swarnakar A,
 PI Tang YT, Warren BA, Yang J, Yue H, Zebardjian Y, Zheng W;
 XX MPI; 2003-403125/38.
 DR N-PSDB; ACC44313.
 XX New human structural and cytoskeleton-associated proteins (SCAP) useful
 PT for diagnosing, treating and preventing diseases or conditions associated
 PT with the aberrant SCAP expression e.g. cancer, osteoporosis, or epilepsy.
 PS Claim 1, Page 258-259; 361pp; English.
 XX This sequence represents a novel isolated human structural and
 CC cytoskeleton-associated protein (SCAP) polypeptide. The polypeptides and
 CC polynucleotides encoding them are useful in diagnosing, treating and
 CC preventing diseases or conditions associated with the decreased
 CC expression or over expression of SCAP, such as cell proliferative (e.g.
 CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's
 CC disease, stroke), heart (e.g. hypertension, heart failure, angina) and
 CC skeletal muscle disorders (e.g. osteoporosis, osteoarthritis) or viral
 CC infections. These are also useful in assessing the effects of exogenous
 CC compounds on the expression of nucleic acid and amino acid sequences of
 CC SCAP. The SCAP or its fragments are useful in screening compounds for
 CC effectiveness as agonist or antagonist of the polypeptides, or in
 CC altering the expression of the target polynucleotide and compounds that
 CC specifically bind to or modulate the activity of the polypeptide. The
 CC microarray is useful in monitoring or measuring protein-protein
 CC interactions, drug-target interactions, and gene expression profiles
 XX

XX Sequence 652 AA;
 SQ Query Match 5.8%; Score 191; DB 6; Length 652;
 Best Local Similarity 19.0%; Pred. No. 5.8e-06;
 Matches 132; Conservative 110; Mismatches 257; Indels 196; Gaps 28;

QY 74 WGGQOQGRGNNVSGRGNNGRGIOANISGRGRLSKRYNNFVAPP--VSR 130
 DB 4 WGGQOQGRGNNVSGRGNNGRGIOANISGRGRLSKRYNNFVAPP--VSR 130
 QY 131 PPL---EGW-----NWQARGSSAQHTVQEPPEVDVDVNASSE 166
 DB 62 PDLNLFKKGMSILDEPEPPSPSLTTTSTSQW---KHWFLVLTDSLSKTYRDSAE 115
 QY 167 EENDSDALDSDDDLASDDVDVQSHSGSRKQNK--WPKKFGGS 210
 DB 116 E---ADELDGELDLRSCDVEYAVQNRVYGFQIHKKAVYLTLSMTSGIRNNWLEAKRT 172
 QY 211 LDSLSTIEGI---NEPORQWHCPACQNGP-----GAIDW 240
 DB 173 VAPTPAPVTKLSQNKENALHSYSTQKGPKAGGGRASGVISRGPRKADGQKQALDY 232
 QY 241 YNLHPL---AHATKARARVKLHRELAIVLEKOLQMGASVIPCSEITYQWKGLGEDEKDYK 283
 DB 233 VELSPLTQASPOKARTPTAPDRLAKO--EELERDLAQRSEERRKWFATDSRTPEVPAG 290

QY 284 ELYGQKGLGEDEKDYEIWPPVIMTLRLDKNDKX-----LGMGN 327
 Db 291 E--GPRGLG-----APLTDQNRRLSEIKKMQLEKLPLRENKVPPLTALN 338
 QY 328 Q-----ELLEYFDKYELRAR-HSYGPGHGMSTLMESSA---TGLLEAE 370
 Db 339 QGRGERGPPSPDGHEALE--KEVQALRAQLLEAWRLQGEAPQSLRSQEDGHIIPPGYISOE 396
 QY 371 RLHRELAEWG-----LDRIAMQKSMFSGVRLQYGLATKODDIFNOHSQGTPLK 424
 Db 397 ACERSLAEHSSHQWMEIQRHHELELR-----LQEKEMULAEETATATSAIE 447
 QY 425 FELKSYQEVNVALEKRLISEDNQOLNYFKKLSKXNK-HAKVLEESLEIWSKEL-RTAAD 482
 Db 448 AMKKAQVEELSELK-----TRSLQGGPGLRKQHSQVDEALKRELQVLSQEQSCLFI 503
 QY 483 NRIYQRTMQH-----EQNREMDAHDRPFMDISIQIHERRDAKENEPMLOQQRKAVY 537
 Db 504 GALMRQAEERETLRRCQOEGELRNH-----OELHGR--LSEETDQLRGFLASOG 553
 QY 538 VGGQQQONINPSSNDCKRAEBVSFIEPOKEMEEFVEREMLIKDOE-----K 587
 Db 554 MNGCCRSNERSCELEVLIRYKANELQYUKKEVQCLRLDELQMMQKDKFTSGKYDQYV 613
 QY 588 KMEDMKRHHHEIFDLEKEFDEALBQIMYKHGLAN 622
 Db 614 ELSHITRSEREIBQLEKHLRLAMALQEKESMRN 648

RESULT 3

AAW81171
 ID AAW81171 standard; protein; 1972 AA.

AAW81171;

05-MAY-1999 (first entry)

Human BAZ2-beta protein.

Transcriptional regulator; transcription; BAZ1-alpha; bromodomain; BAZ;
 atypical zinc finger; testis; human; tumour; BAZ1-beta; BAZ2-alpha; drug;
 BAZ2-beta; treatment; cancer; proliferative disorder; screening.

Homo sapiens.

MO9847920-A1.

29-OCT-1998.

17-APR-1998; 98WO-JP001783.

18-APR-1997; 97JP-00116570.

24-OCT-1997; 97JP-00310027.

(CHUS) CHUGAI RES INST MOLECULAR MEDICINE INC.

Jones MH;

WPI; 1998-583603/49.

DR N-PSDB; AAV68403.

Transcriptional regulator gene family containing bromodomain - may be
 expressed in testis tissue and is useful for treatment of cancer and
 other proliferative disorders.

Claim 1; Page 100-116; 187BP; Japanese.

This sequence represents the human BAZ2-beta protein, a member of a
 family of transcriptional regulator genes containing a bromodomain (BAZ;
 Bromodomain with Atypical Zinc Finger) which are expressed specifically
 in testis tissue and also in certain tumour lines. Transgenic cells may
 be used for the preparation of the BAZ1-alpha, BAZ1-beta, BAZ2-alpha and
 BAZ2-beta proteins. These proteins can be used in the treatment of cancer

CC and other proliferative disorders, and in screening of compounds for
 CC their binding ability to the expression products (e.g. for use as drugs
 CC by modulation of transcriptional regulation)

XX Sequence 1972 AA;

Query Match 5.5%; Score 182; DB 2; Length 1972;
 Best Local Similarity 20.4%; Pred. No. 0.00014;
 Matches 111; Conservative 83; Mismatches 192; Indels 158; Gaps 19;

QY 140 QARGSAGHTANQVEPFDVDDVDNASEENDSDALDDSDDLAS-----183
 Db 394 QRRGTDSDIPSSKSDSESDNEDEDEDEDEDEDESDSDQSDSNSSESDTGSSEE 453
 QY 184 --DDVSDVSQSHSGRKONKWKPFKFGSL--DSLISIQINPQROWHCPACONGGALD 239
 Db 454 DDDDDQDSDSDTDESEKSMKLNKTTSSVKSPPSNLTHSPRNLIHAKESAPALIC 513
 QY 240 WYNLHPL-----AHARTGARVKLHRELAEVLKED-----LQMK 275
 Db 514 SSSQSPAFITGSSSTLTSSPHSGTSKRRVTDRELRILEYGMQRETRIRNPGRLQGE 573
 QY 276 GASVIFPGETIYQW-----KGLGEDEKD-----YEIV-- 303
 Db 574 VAYTAPCGKKLQYPEVIKYLSRNIGIMDISRNFSPAKIRVGDFFEARDGPEQMCCIL 633
 QY 304 -----PPWYIIMNTRLDKNDKMLGMQELLEVEYDKEA-----LRASHVGPQHRG 353
 Db 634 KEEDVIFRIAMEGRGRPPNP-----DRQARBESEHRRRRKGRPPVNDNA 679
 QY 354 MSVLMFESSATGYLLEAERLHRELAEWGLDRIAMQKSMFSGVRLQYGLATKODDIF 413
 Db 680 EFLDNADATLRLKQAEIARQAOIKLRLKQKQEAQVABAQKQQAIMAEE-----734
 QY 414 NQHSQGTFLKTELK-SYEMVVKELRQISEPN---QOLNFKKLSKXNKFAKLEES 468
 Db 735 -----KRQKQKQIKMKQEKIKRIQOITRMEKELELAQOILAKKKKEAANAKLLEAE 788
 QY 469 LEIMSEKLRATMDRIVQRTKMOHEONREEMDAHDRPFMDISIQIHERRDAKENEPM 528
 Db 789 KRIKKEKEMR-----QOAVLKH-QERERRRQH---MMLKAMEARKKAEEK--ER 833
 QY 529 LQOERAKVVGQOQONINPSSNDCKRAEBVSFIEPOKEMEEFVEREMLIKQEK 588
 Db 834 LKQEK-----DEKRLN-----KERTLEQRLELEW-AEKLKRP 866
 QY 589 MEDM 592
 Db 867 NEDM 870

RESULT 4

ABR64241
 ID ABR64241 standard; protein; 1972 AA.

ABR64241;

15-OCT-2003 (first entry)

Angiogenesis protein BNO150.

Cytostatic; antineoplastic; antiarthritic; antidiabetic; ophthalmological;
 KW antipneumatic; antiarteriosclerotic; cardiac; vasotropic; angiogenesis;
 KW gene therapy; vasculature; cancer; rheumatoid arthritis; psoriasis;
 KW diabetic retinopathy; cardiovascular disease; atherosclerosis;
 KW ischemic limb disease; coronary artery disease.

Homo sapiens.

MO2003027285-A1.

03-APR-2003.

19-SEP-2002; 2002MO-AU001282.
 27-SEP-2001; 2001AU-00007973.
 27-SEP-2001; 2001AU-00007974.
 11-OCT-2001; 2001AU-00008210.
 29-OCT-2001; 2001AU-00008532.
 13-NOV-2001; 2001AU-00008838.
 28-AUG-2002; 2002AU-00951032.
 (BION-) BIONOMICS LTD.
 Gamble JR, Hahn CN, Vadas MA;
 WPI; 2003-354655/33.
 N-PSDB; ACF34516.
 New angiogenic genes and polypeptides, useful for diagnosing, prognosticating or treating an angiogenesis-related disorder, e.g. cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or cardiovascular diseases.
 Claim 15; SEQ ID NO 176; 90bp; English.
 The invention relates to the isolation of novel genes (ACF34446-ACF34559) encoding proteins (ABR6180-ABR64281) involved in the process of angiogenesis. The nucleic acid molecules are useful in identifying and/or obtaining full-length human genes involved in an angiogenic process. The nucleic acid molecule, polypeptides or complexes encoded, cells or genetically modified non-human animals derived from these are useful for the screening of candidate pharmaceutical compounds used in treating, angiogenesis-related disorders. They are also useful for diagnosing, prognosticating or treating an angiogenesis-related disorder, which involves uncontrolled or enhanced angiogenesis or is a disorder in which a decreased vasculature is of benefit (e.g. cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or cardiovascular diseases such as atherosclerosis), or involves inappropriately arrested or decreased angiogenesis or is a disorder in which an expanding vasculature is of benefit (e.g. ischemic limb disease or coronary artery disease). The modulator of expression or activity of the polypeptide encoded by the nucleic acid sequence is useful for manufacturing a medicament for the treatment of an angiogenesis-related disorder. This sequence corresponds to one of the novel angiogenic protein
 Sequence 1972 AA;
 Query Match 5.5%; Score 182; DB 6; Length 1972;
 Best Local Similarity 20.8%; Pred. No. 0.00014;
 Matches 113; Conservative 83; Mismatches 190; Indels 158; Gaps 20;
 140 QARGSAGHTAVQEPFVEDVDVNAEEENDSDALDDSDDLAS-----183
 394 QFGCTGSDLPSSKSDSDNEDEDEDEDEDEDEDEDEDSQSSDSNSSEDTGSEEE 453
 184 --DDYSDVSQKSHGRKQKMKFFGSL--DSLSIEQINPQKQWCPACONGPAID 239
 454 DDDDKQODESDTEGSEKTSMLKNTTSVSKSPSMULTGHSPTNMLIAKAPGAPALC 513
 240 WYNIHLPL-----AHARTKAGRRVKRLRLAEVLEK-----LQMR 275
 514 SEQSQPAFLQTSSTLTSFSGTSKRRRYTDERELRILEYQWQRETRINRFGRLQGE 573
 276 GASVTPGGEIYQW-----KGLGDEKQ-----YEIVW---303
 574 VAYVAPCGKLRQVPEYIKLRSNGIMDISRDNFSAKIRVDFEYARDGQEWQWCLL 633
 304 -----PPWVIMNTRLDKDNKWLGMNGQELLEFDKY--EALPAHSYEPQGRGMS 355
 634 KEDEVIRIRAMERGRPRP-----DRQPARRESMRRKGRPPVGVNA 679
 356 VLMFESSA--TGYIEARLRLAEMLGDLRIAMQKRSMSGGVROLYGLATKQDLIDF 413
 680 EFLNADAKLRLKLOAQOEIARQAAQIKLRLKQOEQARVAKAKQQAIMAEF-----734

414 NQHSQGRLLPELK-SYQENVVKELRQISDN-----QQLNYFNKLSKONKAKVLEES 468
 735 -----KKQKEQIKIMKQOEKIRIQIQRWEKLRPAQOILFAKKKKKEAANKLLEAE 788
 469 LIMESEKLRRTAEDNRIRVORTKQHEQNRREMDAHRFFMDSIKQIHERDRAKEENFEM 528
 789 KETKEKEMR-----QQAVLLKH-QERERRQH-----MLMKAMEARKKAEEK--ER 833
 529 LQOGERAKVVGQOQONINPSSNDDCRKAEEVSSSTIEQEKEMEFVEERMLIKDOEKK 588
 834 LKQEKR-----DEKRLN-----KERTLEORLELEM-AKELKKP 866
 589 MEDM 592
 867 NEDM 870
 RESULT 5
 AAE32117
 ID AAE32117 standard; protein; 615 AA.
 AAE32117;
 24-MAR-2003 (first entry)
 Human cytoskeleton-associated protein, CSAP-15.
 Human, cytoskeleton-associated protein, CSAP-15; atherosclerosis; cancer; gene therapy.
 Homo sapiens.
 Key Location/Qualifiers
 FT Domain 66..174 /note="pH domain"
 WO200279404-A2.
 10-OCT-2002.
 25-MAR-2002; 2002MO-US009288.
 29-MAR-2001; 2001US-0280508P.
 03-APR-2001; 2001US-0281323P.
 13-APR-2001; 2001US-0283769P.
 04-MAY-2001; 2001US-0288609P.
 10-MAY-2001; 2001US-0290518P.
 18-MAY-2001; 2001US-0291870P.
 29-MAY-2001; 2001US-0294451P.
 (INCY-) INCYTE GENOMICS INC.
 Hafajia AJA, Tang TY, Yue H, Khan FA, Ison CH, Baughn MR;
 Warren BA, Duggan BM, Thangavelu K, Honnell CD, Azimzai Y;
 Elliott VS, Burford N, Ding L, Yue H, Becha S, Emerling BM;
 Richardson TW, Lee SY, Bandman O, Lal PG, Lee S, Gietzen KJ;
 Walla NK, Griffin JA, Lee BA, Swarnakar A, Ring HZ, Jones KA;
 WPI; 2003-092894/08.
 N-PSDB; AAD49604.
 New human cytoskeleton-associated proteins, useful for preparing a composition for diagnosing or treating a disease or condition associated with decreased expression or overexpression of functional CSAP e.g., cancer.
 Claim 1; Page 177-178; 233pp; English.
 The invention relates to new human cytoskeleton-associated protein (CSAP) and its polynucleotide. The polypeptide is useful for preparing a composition for diagnosing or treating a disease or condition associated with decreased expression or overexpression of functional CSAP e.g., atherosclerosis or cancer. The present sequence is human CSAP-15 protein.

QY 554 RRAEVSFFIEPQEKMEFVEREMLIKDEK-----KMEKMKRHHIEFL- 603
DB 929 RSIABQYSDLE-KEKIMKE-LEIKEMARHKEITKDATIASLETNRLTSDVANLA 986
QY 604 -EKE-----FDEALEQL 614
DB 987 NEKEELNKLKEAQEQL 1003

RESULT 7

AAM79264
ID AAM79264 standard; protein; 1203 AA.

AC AAM79264;
XX 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 1926.
XX
XX

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KM tissue growth factor; immunomodulatory; cancer; leukaemia;
KM nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX MO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001MO-US004098.

XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y,
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AB, Yang Y, Wejthman T, Goodrich R;

XX WPI; 2001-476283/51.
XX N-PSDB; AAK52397.

PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.

XX Claim 20; Page 4327-4329; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX sequence listing were missing at the time of publication

XX Sequence 1203 AA;

Query Match 5.4%; Score 179; DB 4; Length 1203;
Best Local Similarity 19.0%; Pred. No. 0.00012;

Matches 173; Conservative 121; Mismatches 284; Indels 332; Gaps 40;
QY 15 OGGRPE-----YEQVYGLAGTR--LASSQDDGGEWEVSKKNKPKNTSGKTW 63
DB 37 RGRPRPADARASTYGAIVVQGIAGOPFVVLNSGEGGDSFGVQIGANDOG-ASGA-- 93
QY VSONSNPPRAWGGOOQGRGNSVSGRGNVSGRGNNGRI-----QANISGRGRLSRKYD 119
DB 64 LSSDLELPENYSQVKFPAPSQSSTDSBEPGAYWNGKILRSHSQASLAGPVPDSNRS 153
QY 94 LSSDLELPENYSQVKFPAPSQSSTDSBEPGAYWNGKILRSHSQASLAGPVPDSNRS 153
DB 120 NNFV-APPVSRP-----PLE-----GGMWQARGSGAQTAVQEPFVVE 158
QY 154 NSMLETAPKVASPGSTIDTAPLSSVDSLINFDSQLG--QARGRTGRTRM--LPEEQ 208
DB 159 DDVDVASEENDSDALDSDDD-----LASPDYSDVSQKSHGSRKQKWFKEFGSLDS 213
QY 209 RKRKSLDSRLPRDTFEEREROSTMHTSSRTYDNHVTSTKQPAQSQN-----LSP 259
DB 214 LS-----IEQINPQRMHCPCACN-----GPGALDWY--NLH 244
QY 260 LSGFRSROTQDWVLQSFEPFRSAQDPTMLQFKSTPDLRDQEAAPGSDVDMKATY 319
DB 245 PLLAARTKGAARVYTLHLEAVELEKDLQMRGASVIFG--LYGQ-----WKG 291
QY 320 GILREGSSESTSVR--RKVSIVLEK--WQPLVWVSGSTKAVAGGELTRKYBELORK 374
DB 292 LGEDEKDYELVWPVWIIIMNTFLDKDNDKWLGMNGOELLE-----Y 333
QY 375 LDEEVKTKQKLEPSQVGLEROLEEKTRECSRL--QELLEKRGKGAQSNKELQNMRL 430
DB 334 PDKYEALR-----ARHSYGFQ----- 349
QY 431 LDQGEDLRHGLETVQWELQNKLVHQGPBPAKEVLLKDLLETRELLEVEYBEGKQVEBQL 490
DB 350 --GHRGMSVL--MESSAT-----GYLEAR-- 371
QY 491 RLREBELTALKALAEVVASRDQVEHYRQYQRTQELRSMQATODHVALEBRQM 550
DB 372 -----LHRELAEMGLDRIAMGOKRSMFSGVROLYGFLATKODI----- 410
QY 551 SALVYGLQLELEETSEETGRW--QSMFOKNKEDL--RATKQELLQLRMEKEEMEELG 604
DB 411 --DIFNQH-----SQGKTR--LKPELKYQEMVYEL----- 438
QY 605 EKIEVLQRELEQARASAGTRQVEVYLKKELLRTQE-ELKEIQAEKQSDVAGRHRDELE 663
DB 439 -----ROISEDNQO-----LNVFKNKLKSKONGKAKVLEESLEIMSEK----- 475
QY 664 KOLAVLRVEADRGRELEBQNLQKTLQQLRODCBESAKKAVALEAVLQGRRAAVET 723
DB 476 -LRTAEDNRIVRQRTKQHQEONREEM-----DAHDFEMDSIKQIHERRDAKEBNFEM 528
QY 724 TLRETEQENDEPRRLILGLEQLKETRGVLVDGEAVEALRLDKLRLAEKQLEBALNA 783
QY 529 LQOQE-----RAKV-----VGQOONINPSSNDCKRRAREVVSFFIEFOE- 568
DB 784 SQREBSLAAAKRALPARLEBAORGARIGQOQTLNRLLEEGRQEVLRGKQALEBQ 843
QY 569 -----KEMEFPVEEREMLIKDEKMEKMKRHHIEFLPEK-DEALEQLM 615
DB 844 KRLIDTVDRLNKLKEIKGDSKQALQQLQALQEDVKEKARREAVDAQOQAWMSAEAK 903
QY 616 YKHGLHNEDD 625
DB 904 TSGGLRLQD 913

RESULT 8

ABG96296
ID ABG96296 standard; protein; 1203 AA.

XX AC ABG96296;

11-DEC-2002 (first entry)
Human ovarian cancer marker OV9.
Human; ovarian cancer; marker; cancer; familial history; brain disorder; central nervous system disorder; bacterial meningitis; viral meningitis; Alzheimer's disease; Parkinson's disease; cerebellar edema; hydrocephalus; brain herniation; inflammation; encephalitis; testicular disorder; noncancerous granulomatous orchitis; connective tissue disorder; heart disorder; ischaemic heart disease; atherosclerosis; neoplasm; histological type; carcinogenic; ovarian cancer marker.
Homo sapiens.
WO200271928-A2.
19-SEP-2002.
14-MAR-2002; 2002WO-US007826.
14-MAR-2001; 2001US-0276025P.
14-MAR-2001; 2001US-0276026P.
10-AUG-2001; 2001US-0311732P.
19-SEP-2001; 2001US-0323580P.
26-SEP-2001; 2001US-0324967P.
26-SEP-2001; 2001US-0325102P.
26-SEP-2001; 2001US-0325149P.
(MILL-) MILLENNIUM PHARM INC.
Morahan JE, Gannavaram M, Hoersch S, Kamathkar S, Kovatis SG, Meyers RE, Morrissett MP, Olandt PJ, Sen A, Vieby PO, Mills GB, Baer RC, Lu K, Schmandt RE, Zhao X, Glatt K; WPI; 2002-723277/78. N-PSDB; ABS763388.
Assessing whether a patient is afflicted with ovarian cancer, useful in assessing the stage or progression of the disease, comprises comparing the expression level of a cancer marker in a sample from a patient and from a non cancer patient.
Disclosure; Page 160-163; 481pp; English.
The present invention relates to a new method for assessing whether a patient is afflicted with ovarian cancer. The method involves comparing the expression level of a marker in a patient sample and the normal level of expression of the marker in a control non-ovarian cancer sample, where the marker is selected from 363 cancer markers described in the specification. The method of the invention is useful in diagnosing or characterising cancer, in detecting the presence of cancer as early as possible, and the recurrence of ovarian cancer. The method may also be of particular use with patients having an enhanced risk of developing ovarian cancer (e.g. patients having a familial history of ovarian cancer). The cancer markers may be used in the management and treatment of e.g. brain and central nervous system disorders (e.g. bacterial and viral meningitis, Alzheimer's disease or Parkinson's disease), brain disorders (e.g. cerebral oedema, hydrocephalus or brain herniations), inflammations (e.g. bacterial or viral meningitis or encephalitis), testicular disorders (e.g. noncancerous granulomatous orchitis), connective tissue disorders, or heart disorders (e.g. ischaemic heart disease or atherosclerosis). The compositions and methods may also be used in assessing the histological type of neoplasm associated with ovarian cancer, monitoring the progression of ovarian cancer, determining whether ovarian cancer has metastasized or is likely to metastasize, selecting a composition for inhibiting ovarian cancer, assessing the ovarian carcinogenic potential of a compound, or inhibiting ovarian cancer or at risk of developing ovarian cancer. The present amino acid sequence represents one of the ovarian cancer markers described in the invention
Sequence 1203 AA;
XX

Query Match	5.4%;	Score 178;	DB 5;	Length 1203;		
Best Local Similarity	19.0%;	Pred. No. 0.00014;				
Matches	173;	Conservative	121;	Mismatches	264;	
			Indels	332;	Gaps	40;

QY	15	QGVYRP-----VEOLVQLAGTR--LASSODDGEMEVISKKNNKRGNTSGKTW	63
DB	37	RGGRPPAKDARASTYGVAVRVQGIAGQFVVLNSGKEGDSFGVQIKANQG--ASGA--	93
QY	64	VSQNSNPRAWGCGOQGRGNSVSGRGNVSGRGNNGRT-----QANISGGRLSRKYD	119
DB	94	LSSLDLEPENYSQVKGFPAPSOGSTDEBPAAWNGILRSHSGLASLAPRPVDPNRS	153
QY	120	NNFV-APPVYRP-----PLE-----GGMNQARQSAQHTAVQCFPVE	158
DB	154	NSMLETAPKVASPOSTIDTAPLSSVDSLINFPSDLGG--QAKRTGRTRM--LPFQ	208
QY	159	DDVNASSEENDSDALDDSDDD-----LASDDYSDVSQKSHGRKKQMKPFPGSLDS	213
DB	209	RKRSKSLDSRLPRPTFERERQSTNTMHTSSTKYNHVTGSKQAPQSQ-----LSP	259
QY	214	LS-----TEQINEPQRMHCPCQN-----GPGALDW--NLH	244
DB	260	LSGSRSRQTDWVLQSFEEPRRAADPTMLQFSTPDLRLDQEEAPRGSVDHMKATLY	319
QY	245	PLAHARTKGRARKYLHRELAEVLKDLQMGAGAVIPGQ---LYGQ-----WKG	291
DB	320	GILGEGSESETSYR--RKSLVLEK--MPLVWSSGSTKAVAGQSELTTRYVELQK	374
QY	292	LGEDEKDEIYMPVVIIMTRLRKDDNDKMYGMGNQELLE-----Y	333
DB	375	LDEEVKKRQKLEPSQVGLEROLEETECESRL-----QELLRRKGEAQQSKNELQNMRL	430
QY	334	FDKYTEALR-----ARSHYGPQ-----	349
DB	431	LDQGEDLRHGLQETQWMLQNKTKVQGPAPAKVLLKDLTRELLEEVLEBQKQVEQL	490
QY	350	--GHRGMSVL--MFESSAT-----GLTEAR---	371
DB	491	RLRRLELTALMGALKEEYASRDOEVEHYRQOYORDTEQLRRSMODATODHAYLAEARKM	550
QY	372	--LHRELAENGLDRIAMQKRSMPSSGVRQYGFATKQDPL-----	410
DB	551	SALVRGLQRELEETSEETGHW--QSMFQKNKEDL--RATKQELLQRLMEKEBEBELG	604
QY	411	--DIFNOH-----SQGKTR--LKFELKSYQEMVVKEL-----	438
DB	605	EKIEVLQRELEQARASAGDTRQVEVLKKELLRTQE--ELKELOAKRQSQEVAGRHRRLE	663
QY	439	--ROISEDNOQ---LNFYKRLSKQNKHAKVLEESLEINSEK-----	475
DB	664	KQLAVLRYEADRGRELBEQNLQLOKTLQOLRQDCEBAKAKMVVAEAVTLVQORAAVET	723
QY	476	--LHRTAEENRIVRQRTKMQHEONREEM-----DAHDFPMDSIKQIHERRDAKEENFEM	528
DB	724	TLRELTQEBENDEFRRRIIGLEQOLKETRGLVDGGEVAVELRLDKQORLEAKQOLEBALNA	783
QY	529	LQOQB-----RAKY-----VGOQOQININSNDCKRAEYVSFTFQF-	568
DB	784	SOEBEGSLAAKRALREARLEAEGCLARLAGEQOTLINALBEGGQREVLRRGKALEBQ	843
QY	569	--KEMEEFEEEREMILIKQEKKEMEDMKGRHHEIFDLKEF--DEALQOLM	615
DB	844	KRLLDRTVDRINKLEKIGEDSKQALQLOLQALQLEDYKEKARREVAADQARQADWASAEK	903
QY	616	YKHGLHNEDD	625
DB	904	TSGLSRLQD	913

RESULT 9	
AAAM25750	
ID	AAAM25750 standard; protein, 1879 AA.
XX	

AA025750, 16-OCT-2001 (first entry)
Human protein sequence SEQ ID NO:1265.
Human, cancer; ulcer; HIV infection; human immunodeficiency virus; antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; vitruclide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia; antiagregant; haemostatic; vulnary; antilcer; osteopathic; eczema; dermatological; antiallergic; antiasthmatic; antidiabetic; cystostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder.
Homo sapiens.
MO000153455-A2.
26-JUL-2001.
22-DEC-2000; 2000WO-US035017.
23-DEC-1999; 99US-00471275.
21-JAN-2000; 2000US-00488725.
25-APR-2000; 2000US-00552317.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Drmanac RT;
WPI; 2001-457603/49.
N-PSDB; AA099691.
Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
Claim 20; Page 262; 1217pp; English.
AA099166 to AA099904 encode the human proteins given in AA025225 to AA025963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; vitruclide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antiagregant; haemostatic; vulnary; antilcer; osteopathic; dermatological; antiallergic; antiasthmatic; antidiabetic; cystostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders
Sequence 1879 AA;
Query Match 5.3%; Score 175; DB 4; Length 1879;
Best Local Similarity 20.2%; Pred. No. 0.00046;
Matches 116; Conservative 94; Mismatches 194; Indels 170; Gaps 24

QY	158	EDVDVDA	SE-----	-----	ENDSALDSDSDLLASDVDSVDSQSHSRKONKWEFK	206				
		::: :	::: :	::: :	::: :					
Db	1297	QKEIENITQ	YEEKAAAYDKLETKTKRLQOELDVLV	-DLDNQRLVSNEKKQRK	-FDQ	1354				
QY	207	PFGLSDLSL	EQINEPQRQWCHPCAQNGPAIDWYNLHPLLAHRTGABRVKLHRELA	E	266					
		::: :	::: :	VEABEREKETALSLARALE	1393					
Db	1355	LIAEKNIS	SKVADERR-----							
QY	267	VLEKDLQ	MGSAVIPCGEI-----	YGOMKGLGED	-EKDYELIYMPVVIIMTRLD	315				
		::: :	::: :	::: :	::: :					
Db	1394	ALFAKEKE	LERTNTKMLKAEMGRPGSASKDVQO	-ELSHDLEKSKRALGDPRLEEMKTQLE		1451				
QY	316	KDNDKML	GMGNQEL-----	LEY-----	FDKYELARAHSYGPOSHRGMSV	356				
		::: :	::: :	::: :	::: :					
Db	1452	E-----	IGRTLELASPRDDAKIRLEVMQAPSRASFER	-DLQATTEQNEERRLQR		1501				
QY	357	IMFESSAT	GYLEARELHRELAEMGLDRIAWGQKRS	-----	FSGG--VRQLYGFLA	405				
		::: :	::: :	::: :	::: :					
Db	1502	QIHHEYTE	--LEDERKQRALAAAKIKLGMDPVYTTLDLXADSATIKRGGAIKQRLRQ	A	1559					
QY	406	TKQDL	-----	DIF---NQSHQGKTRLFKELKSQYEMVVKELRQISENQOLNY		450				
		::: :	::: :	::: :	::: :					
Db	1560	QMKDFQ	RELDARASRDEIFATAKENKKAKSLIADLMQLOE	-----	DLAAABGRKQADL	1615				
QY	451	FKNKL	-----	KONKAKV--LEESL-----	EIMSEKLR-----	478				
		::: :	::: :	::: :	::: :					
Db	1616	EKEHLAE	ELASSLSGRNALODEKRLERLARIQLEEELEEBEGQNEAMSDRPRKATQQA	AEQ		1675				
QY	479	-----	TAEDNRIVRQRTKQWQHNEBMDAHDPRFFWDSTIKQIHERRDAK	---	EE	524				
		::: :	::: :	::: :	::: :					
Db	1676	LSNELATER	STAQKNESARPOOLEQNKELSKLHEMGAVAKSKFSTYLALEAKIAQLE			1735				
QY	525	NFEMLQ	QOEBAKAVGQGOQONINPESNDCKRRAEVSFTIEFOKEMEEPFEKE	-----		579				
		::: :	::: :	::: :	::: :					
Db	1736	QVBEA	ERKQATKSLKQKD-----	KGLKEILLQYEDERKMAEQYKQAEKGNAR		1785				
QY	580	-MLIKD	QEKTMEDMKRHHBEIPDLKEPDEALE			612				
		::: :	::: :	::: :	::: :					
Db	1786	VKQAKR	QLEAEESQRIINARRQLQGELEDATE			1819				
RESULT 10										
ABRS7420										
ID	ABRS7420	standard; protein; 580 AA.								
XX	AC	ABRS7420;								
XX	DT	15-SEP-2003 (first entry)								
XX	DE	Human NOV4b protein SEQ ID NO:18.								
XX	XX									
KM	Human; NOVX; cytosolic; cardiac; antiinflammatory; immunosuppressive;									
KM	antiallergic; hemostatic; anti-HIV; antidiabetic; antiarteriosclerotic;									
KM	anorectic; antistomatitic; nephrotropic; antiarthritic; hepatotropic;									
KM	neuroprotective; nootropic; antidiabetic; vitruclide; antiparastolic;									
KM	relaxant; anticonvulsant; hypotensive; vasotropic; antiparkinsonian;									
KM	vulnerary; angiogenic; antiangiogenic; gene therapy; vaccine; cancer;									
KM	cardiomyopathy; atherosclerosis; hypertension; diabetes; inflammation;									
KM	autoimmune disorder; allergy; blood disorder; AIDS; obesity; asthma;									
KM	acquired immunodeficiency syndrome; nephropathy; cirrhosis; arthritis;									
KM	Alzheimer's disease; Parkinson's disease; goitre; infection; stroke;									
XX	muscular dystrophy; epilepsy; wasting disorder; chromosome 5.									
XX	OS	Homo sapiens.								
XX	XX	WO200294870-A2.								
XX	PD	28-NOV-2002.								
XX	PF	02-NOV-2001; 2001WO-US051580.								
XX	PR	02-NOV-2000; 2000US-0245291P.								

PR 02-NOV-2000; 2000US-0245317P.
 PR 07-NOV-2000; 2000US-0246562P.
 PR 08-NOV-2000; 2000US-0246871P.
 PR 26-JAN-2001; 2001US-0264389P.
 PR 26-JAN-2001; 2001US-0264423P.
 PR 29-JAN-2001; 2001US-0264799P.
 XX
 PA (CURA-) CURAGEN CORP.
 PI Grosse WM, MacDougall JR, Smithson G, Millet I, Stone DJ;
 PI Gunther E, Ellerman K, Alsobrook JP, Lepley DM, Burgess CE,
 PI Sytek KA, Edinger SR, Gangoli EA, Gorman L, Taupier RJ, Li L,
 PI Guo X, Fernandes BR, Vernet CM, Tcherven VT, Casman SD, Shenoy S;
 PI Mishra V, Furtak K, Baumgartner JC, Colman SD;
 XX
 DR WPI; 2003-140359/13.
 DR N-PSDB; ACF03355.
 XX
 PT New NOVX polypeptide useful for preventing or treating NOVX-associated
 PT disorders, e.g. cancer, cardiomyopathy, atherosclerosis or diabetes, and
 PT in chromosome mapping, tissue typing or pharmacogenomics.
 XX
 PS Claim 1, Page 45; 346pp; English.
 XX
 CC ACF03547 to ACF03570 encode the human NOVX proteins (I) given in ABR57412
 CC to ABR57435. (I) have cytoskeletal, cardiac, anti-inflammatory, nootropic,
 CC immunosuppressive, antiallergic, haemostatic, anti-HIV, antidiabetic,
 CC antitumor, antitoxic, anorectic, antisthmatic, nephrotropic, virocidic,
 CC antiparasitic, hepatocytotoxic, neuroprotective, antibacterial, relaxant,
 CC antiparalytic, anticonvulsant, hypotensive, vasotropic, antiparkinsonian,
 CC vulnary, angiogenic and antiangiogenic activities, and can be used in
 CC gene therapy and vaccines. The NOVX polypeptides and their antibodies can
 CC be used to determine the presence or absence of (I) in a sample. The NOVX
 CC polypeptides, polynucleotides encoding them, and antibodies against them,
 CC are useful in manufacturing a medicament for treating or preventing a
 CC syndrome associated with a NOVX-associated disorder such as hypertension,
 CC cardiomyopathy, atherosclerosis, cancer, diabetes, asthma, inflammation,
 CC autoimmune disorders, allergies, blood disorders, obesity, acquired
 CC immunodeficiency syndrome (AIDS), immunoglobulin (Ig)A nephropathy,
 CC cirrhosis, arthritis, Alzheimer's disease, Parkinson's disease, goitre,
 CC infections (e.g. bacterial, viral, parasitic), stroke, muscular
 CC dystrophy, epilepsy, and other wasting disorders associated with chronic
 CC diseases. ACF03571 to ACF03644 represent PCR primers and probes for NOVX
 CC sequence, which are used in an example from the present invention
 XX
 SQ Sequence 580 AA.

Db 322 SOBACERSLAEMSSHQVMEIQRHHERELOR-----LQOEKMWLAETATAS 372
 QY 422 RLKPELXSVQEMVVKELRQISEDNQOLNFKNLKSKONK-HAVYLESLIEMSEKL-RT 479
 Db 373 ALEAMKKAQOEISRELSK-----TRSLQOCPPDGLRKQHSQDVALKRELOVLSQK 428
 QY 480 AEDNRIVRQTKNQH-----EQNREEMDAHDFEFMSIKOIHBRPAKENFEMLOQOER 534
 Db 429 LEIGALMRQAEERHEHTLRQOQEGELLRN-----QELHGR-----LSEIDQLNGFIA 478
 QY 535 AKVAGQOQONINPSSDDCKRAEVSFLFEQEKMEFEVEREMLIKQF-----586
 Db 479 SQGWGCGGRSNRSESCVELLVLRKVENELQYLKKEVQCIRDELQMWQKDKRFTSGKYOD 538
 QY 587 --KKMEDMKRHHHEIFDLEKEFEDEALQOLMYKHGLN 622
 Db 539 VYVELSHIKTRSERETQLKEHLRLMAALQEKSEKMN 576
 XX
 PT RESULT 11
 PT ABP62858
 ID ABP62858 standard; protein; 593 AA.
 XX
 AC ABP62858;
 XX
 DT 14-OCT-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 295.
 XX
 KW Human; vulnary; dermatological; neuroprotective; nootropic; cancer;
 KW antiParkinsonian; immunostimulant; cytoskeletal; immunosuppressive;
 KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
 KW burn; central nervous system disorder; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; immune disorder;
 KW autoimmune disease; multiple sclerosis; diabetes; allergy.
 XX
 OS Homo sapiens.
 XX
 PN WO200218424-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 31-AUG-2001; 2001WO-US027093.
 XX
 PR 01-SEP-2000; 2000US-00654935.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Aunndi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
 PI Zhao QA, Wang D, Liu C, Dymnac RT, Wehrman T;
 XX
 DR WPI; 2002-583321/62.
 DR N-PSDB; ABQ93337.
 XX
 PT New polynucleotide and polypeptides, useful for treatment and diagnosis
 PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
 PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
 PT sclerosis, diabetes and allergies.
 XX
 PS Claim 20; SEQ ID NO 295; 284bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising one of
 CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising
 CC administering to a mammalian subject a composition comprising the protein
 CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
 CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
 CC (I) is useful for gene therapy of diseases and (II) can be used for
 CC therapeutic treatment. Diseases that may be treated include wound healing
 CC and tissue repair, burns, central nervous system disorders (e.g.
 CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
 CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
 CC sclerosis, diabetes and allergies. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 593 AA;

Query Match	5.2%;	Score 172.5;	DB 5;	Length 593;
Best Local Similarity	19.0%;	Pred. No. 0.00014;		
Matches 110;	Conservative 100;	Mismatches 209;	Indels 161;	Gaps 23;

```

OY      162 DNASEENDSDALDDSDDDDLASDDYDSDVSQKSHSGSRQNK-----WFK 205
Db      52 DSTAEE---ABDLGEIDLRCSTDTVEAVQANVGFQJHTKDAVYTLASMTSGIRRWIE 108
OY      206 KFGSSLSLSIEOI-----NEPQRCWCPACONGP----- 235
Db      109 ALRKTVPRTSADPYTKLSDSNKENALHSYSTQKGLPKAGEORAGSEVI SRGGPRKADQR 168
OY      236 GAIDWYNLHPLL-----AHARTKAGARVYLHRELAVLEKDLQMRG-----AS 278
Db      169 QALYVEISPLTQASPPQARFPARTPPDLAKO--BELERDLAKRSEBRKXWFEATDSRT 226
OY      279 VIPGGEIYQWKGGLGEDEKDEYIWPMPWIIIMNTLBDKDNDK----- 322
Db      227 EVPAGE--GPRRGGL-----APLTEDQNNLSIEIKKQWLEKTLPLRENKRVPL 274
OY      323 LGMGNQ-----ELLEFDKYEALRAR--HSYGPQGRGWSVIMPESSA---TG 365
Db      275 TALNQSNGRGERGPPSDGHEALE--KEVQALRAQLLEAWRLQGEAPQASALRSQEDGHIPRG 332
OY      366 YLEAERLHRELAEWG-----LDRITAMGQKRSMPFGGVQOLYGFPLATKDDLDIFNQHSG 419
Db      333 YISQACERSLAEWESSHQUWWEIQRHHERELQR-----LQOEKEMLLAESTAT 383
OY      420 KTRLKFELKSVOEWYKELROI SEDNOOLATPYKNIKLSQNK--HAKYLESEIIMSEKL-R 477
Db      384 ASATIAMKKAAYEEISRELSK---TRLSQGGPDGIRKQHOSDVYALKRELQVLSBSQYQ 439
OY      478 RTAEENRIYRQRTKMOH-----EONKEENDADHDFPMSIKOIHERDAKENFEMLQOO 532
Db      440 KLELIGALMRQAEERHRTLARCQOBSQELLRN-----QELHGR--LSEIIDLRGF 489
OY      533 EBAKVVGQOOQONINPSSNDCKRKRAEVSFIEFOEKEMEFYEEHEMLIKOE----- 586
Db      490 IASQGGNGCGSNNERSCELEVTLRVKNENELQYLKEVQCLDELQOMQOKDRFTSGKY 549
OY      587 -----KKNEDMKKKRHHHEIFDLKEPFDEALEQMLYKGLN 622
Db      550 QDVVYELSHIKTRSERIEOLKHNHLALMAALQEKSMEN 569

```

RESULT 12	
ABG08600	
ID	ABG08600 standard; protein; 923 AA.
XX	
AC	ABG08600;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #8591.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
PR	

XX (HYSE-) HYSEQ INC.
PA
XX
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS72787.
XX

aa New isolated polynucleotide and encoded polypeptides, useful in
 pr diagnostics, forensics, gene mapping, identification of mutations
 pr responsible for genetic disorders or other traits and to assess
 pr biodiversity.

PS Claim 20; SEQ ID NO 38959; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders of altered expression of (II) or biological activity. The invention involves aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG03177 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/publshd_pct_sequences

Sequence 923 AA;

Query Match	5.2%;	Score 172.5;	DB 4;	Length 923;
Best Local Similarity	19.7%;	Pred. No. 0.00027;		
Matches 60; Conservative	64;	Mismatches 112;		Indels 69; Gaps 9

```

QY      366 EMBRLHRELAEMLGDLRIAMGCRSRSPFSGGVRQOLYGLATKODLDIPNOSOGTRLRKFE- 426
      | : | | : |
Dd      153 EYHNLKRELESYG-----RQLOAEVYNNOMSLNRRQOEHLRQOEER 195
      | : | | : |
QY      427 -----LKSQYEMVVKELRQISEDNQOLNFKNLKSKONKHAKVL 465
      | : | | : |
Dd      196 LMBQERQOEODRLHEQERLREQERLRCOEERLRCOEERLRBHHEERLCEQERLREQ 255
      | : | | : |
QY      466 EESLEIMSEKLRRTAE-----DNRIVRQTKQOHEONR-----EEMDADRPF 508
      | | | | : |
Dd      256 EERLCEQERLHEHEERLCEQERLCEQERLHEQERLREQERLRCOEERLREHEERL 315
      | | | | : |
QY      509 MDSIKOIHERRDAKENFEMLOQOEAKYVVGQOQONINPSSNDDCKR-----AEV 566
      | : | | : |
Dd      316 CEQERLREHEERLCEQERLREQERL--LCEQERLREQERLCEQERLPGQERLLEEV 373
      | : | | : |
QY      561 SSIERQOEK-EMERFVBEREMLIKQOEKME-DMKRKRHHEIFDLKXFDEALQOLMYKH 618
      | : | | : |
Dd      374 EKLLEERRRQOEERLLEERLLEEVKLLQOEQOEERQOEHEMLLEERLLEE-VEKLLQOE 432
      | : | | : |
QY      619 GLHNE 623
      | | |
Dd      433 RRHEE 437

```

RESULT 13
ADB79862
ID ADB79862 standard; protein; 892 AA
XX
AC ADB79862;

XX 04-DEC-2003 (first entry)
 DT Rat myosin heavy chain, SEQ ID 102.
 DE Analgesic; pain; streptozocin-induced diabetes; rat.
 XX Rattus norvegicus.
 OS
 XX EP1279744-A2.
 PN
 XX 29-JAN-2003.
 PD
 XX 26-JUL-2002; 2002EP-00255249.
 PF
 XX 27-JUL-2001; 2001GB-00018354.
 PR 07-FEB-2002; 2002GB-00002910.
 XX (WARN) WARNER LAMBERT CO.
 PA Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
 PI WPI; 2003-395407/38.
 DR N-PSDB; ADB79863.
 XX
 XX Use of isolated gene sequences and encoded polypeptides that are
 PT upregulated in the spinal cord in response to streptozocin-induced
 PT diabetes for screening compounds for the treatment of pain, or for
 PT diagnosing pain.
 CC
 XX Claim 1; Page 181-183; 334pp; English.
 XX
 CC The present invention relates to nucleotide sequences which are useful in
 CC the screening of compounds for the treatment of pain, or for the
 CC diagnosis of pain. The nucleotide sequences are up-regulated in the
 CC spinal cord in response to streptozocin-induced diabetes. The present
 CC sequence was used to illustrate the invention.
 XX
 XX Sequence 892 AA:
 SQ
 Query Match 5.2%; Score 172; DB 7; Length 892;
 Best Local Similarity 21.0%; Pred. No. 0.00028;
 Matches 122; Conservative 98; Mismatches 186; Indels 174; Gaps 29;
 QY 141 ARGGSAQHTAVQFPP-DVEDVDNMASEENDSDALDDSD--DDLASDDYSDVSQKSHGS 197
 DB 259 SRKQAFPTQOIEELKRLQEEVAKASAL--AAALQSSRHDCDLRQYESEQAKAEIQ 315
 QY 198 RKONK-----WPKKFGSLDSLSTIEQINEPQRMWHCPACQNGPAIDWYNLHPLLAHA 250
 DB 316 RAMSKANSEVAQWRITKY-----ETDAIQRTTE-----LEEA 346
 QY 251 RTYGARVKLHRELAEVLEKDLQMGASVIPCSEIYGQWKGIGEDKDYEIWPPWITM 310
 DB 347 KKTLAQRLODAEBEHAIVN-----AKCASLEKTKQRLQNEVEDL-----MIDVE 390
 QY 311 NNR-----LDKDNNDKMLGNGNELLEYPKYVALARHSYGGHGRGKSVLMF----- 359
 DB 391 RTVAACALDKKORN-----FDKILAWKQKYETTHAEIASQKESRSSTELFKIKNAV 445
 QY 360 -----ESSATGYLAEERLHRELAEMLDRIAMGQKR-----SMFGGVAVQLGFL-ATKQDL 410
 DB 446 EESLDQLETLTKRNNKMLQCEISPL-TEQIALGSKRIHELEKIKQIQEHSSELQALAEHA 504
 QY 411 DIFNQSQGK-TLKTELKSYQEMVAVKEQLQISEDNOQLYFNKKLSKONKAKVLEESTL 469
 DB 505 EASLEHEEGKILRIQLEINQVKEISDRKIAEKDEIDQL-----KRN-HIRVYESMQ 555
 QY 470 EIMSELRRAEENRIVRQRTKMGHQNRREMDAH--DRFMDSIK-----QI 515
 DB 556 STIDAEIR--SRNDAIRIKKRGEGDNEWEIQUNHSNRPAALRNRYNTQGIKLDQTOL 612
 QY 516 H-----ERR-----DAK-EENFEMLQOQERAKVVGQOO----- 542

DB 613 HLDLALRGQEDLKEQLAMVERRANLILQAEIEELRATLQETGERSRKIAEQELDASERVOL 672
 QY 543 ---QINPSSNDDCRGAEEVSSFFIFQEKEMEEFVE-----REML 581
 DB 673 LHTQN---TSLINTKRLKETDISQI---QGEEDIVQEARNAEKAKAITDAAMAAEEL 726
 QY 582 IKDOEKK--MEDWKKRHHIEIFDLKEKPEALQMYKKG 619
 DB 727 KKEQDTSALIERMKKMLBQTVKDLQHRLEDA-EQLALKGG 765
 RESULT 14
 ADD47859
 ID ADD47859 standard; protein; 892 AA.
 XX
 XX ADD47859;
 AC
 XX 29-JAN-2004 (first entry)
 DT
 XX Rat Protein AAB29713, SEQ ID NO 13555.
 DE
 XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 XX Rattus norvegicus.
 OS
 XX WO2003016475-A2.
 PN
 XX 27-FEB-2003.
 PD
 XX 14-AUG-2002; 2002WO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-033347P.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PI
 XX Woolf C, D'urso D, Belfort K, Costigan M;
 XX WPI; 2003-268312/26.
 DR GENBANK; AAB29713.
 XX
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PT
 XX Claim 1; Page; 1017pp; English.
 PS
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 892 AA;

Query Match 5.2%; Score 172; DB 7; Length 892;

Best Local Similarity 21.0%; Pred. No. 0.00028;
 Matches 122; Conservative 98; Mismatches 186; Indels 174; Gaps 29;

141 ARGSSAQTAVOFP-DVEDVDNASEENDSDALDSDS-DDDLASDDYSDVSGKSHGS 197
 259 SRGKQAFVTOQIEHLKQLEBEVAKSAL-AHALQSSRHDCDLREQYEEQEKAKELQ 315
 198 RKONK-----WFKRFGSLDSLSTIQINEPQRQWHPACQNGPAGIDWYNLHPLLAHA 250
 316 RAMSKANSEVAOWTKY-----ETDAIQRTEE-----LEBA 346
 251 RTKGARRVKLHRELAEVLEKDLQMRGASVIPCSEIYQMGKLGEBDEKDYEIWPPVITM 310
 347 KKCLAQRLQDAEHEVAVN-----AKCASLKTQRQLQNEVEDL-----MIDVE 390
 311 NTR-----LDKDDNDKMLGNGNOELLEYPDKYALRARSYSGQGRGMSVLMF----- 359
 391 RTNACALDCKQKN-----FDKILAEWKQCYETHALEASQKESKSLSTELFKIKNAY 445
 360 -----ESSATGYLAEARLHRELAEWGLDRIAMGQR-----SMFGGVRLYGLF-ATKODL 410
 446 EESLDQLETLKREKNLQOEISDL-TEQIAGGRIHLEIKIKQIQEKSSELQALAEFA 504
 411 DIFNHSQOK-TRLFELKSYQEWVAVKELQISEDNQNLVFNKLSKONKHAVLEESL 469
 505 EASLEHSGKILRIQELINQVASEIDRKIAEKDEBIDQ-----KRN-HIRVESMQ 555
 470 EIMSEKRLRTAADRIVQRTMGHONREEMDAH--DRFPMDSIK-----QI 515
 556 STLDIEIR---SRDAIRIKKKMEGDLNEMELQLNHSRMAAEALRANRNOGILKDTQL 612
 516 H-----ERR-----DAK-EEHFEMLOOQERAKVVGQQ----- 542
 613 HUDDALRGEDLKEQLAWERRANLLQAEIELRATLEQTSRSKIAEOELLDASERVOL 672
 543 ---QINPSSNDCKRRAEVSFIEFOEKEMEVEER-----REML 581
 673 LHTQN---TSLINTKKQJLTDISQI---QGEWEDIVQEARNAEEKAKAIIIDAMMABEL 726
 582 IKDOEKK-MEDMKKKHHEIFDLKEKFEDEALQOLMYKHG 619
 727 KKEQDTSALHERKKQLLEQTVADLQHRUDEA-EQLALKGG 765

RESULT 15

ID ADD47855 standard; protein; 892 AA.

AC ADD47855;

XX 29-JAN-2004 (first entry)

DE Rat Protein AAB29713, SEQ ID NO 13551.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW chronic constriction injury; CCI; spared nerve injury; SNi; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P;
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-033347P.

PA (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.

PI Woolf C, D'Urso D, Befort K, Costigan M;

XX WPI; 2003-266312/26.
 DR GENBANK; AAB29713.

XX New composition comprising two or more isolated polypeptides, useful for
 DR preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNi)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 892 AA;

Query Match 5.2%; Score 172; DB 7; Length 892;

Best Local Similarity 21.0%; Pred. No. 0.00028;
 Matches 122; Conservative 98; Mismatches 186; Indels 174; Gaps 29;

141 ARGSSAQTAVOFP-DVEDVDNASEENDSDALDSDS-DDDLASDDYSDVSGKSHGS 197
 259 SRGKQAFVTOQIEHLKQLEBEVAKSAL-AHALQSSRHDCDLREQYEEQEKAKELQ 315
 198 RKONK-----WFKRFGSLDSLSTIQINEPQRQWHPACQNGPAGIDWYNLHPLLAHA 250
 316 RAMSKANSEVAOWTKY-----ETDAIQRTEE-----LEBA 346
 251 RTKGARRVKLHRELAEVLEKDLQMRGASVIPCSEIYQMGKLGEBDEKDYEIWPPVITM 310
 347 KKCLAQRLQDAEHEVAVN-----AKCASLKTQRQLQNEVEDL-----MIDVE 390
 311 NTR-----LDKDDNDKMLGNGNOELLEYPDKYALRARSYSGQGRGMSVLMF----- 359
 391 RTNACALDCKQKN-----FDKILAEWKQCYETHALEASQKESKSLSTELFKIKNAY 445
 360 -----ESSATGYLAEARLHRELAEWGLDRIAMGQR-----SMFGGVRLYGLF-ATKODL 410
 446 EESLDQLETLKREKNLQOEISDL-TEQIAGGRIHLEIKIKQIQEKSSELQALAEFA 504
 411 DIFNHSQOK-TRLFELKSYQEWVAVKELQISEDNQNLVFNKLSKONKHAVLEESL 469

```

Db      505 EASLEHEEGKILRIQLELNOVKSEIDRKIAEKDEIDOL-----KRN-HIRVESMQ 555
Qy      470 EIMSEKLRRTABDNRIVRORTKQOEONREEMDAH--DRPFMSIK-----OI 515
Db      556 STDAEIR---SRNDAIRIKKKMEGDLNEMEIQLNHSNMAAEALRNYRNTOGILKDTOL 612
Qy      516 H-----ERR---DAK-BENFEMLOQOERAKVGOQO-----542
Db      613 HLDALRGQEDLKEQLAMVERRANLQAEIBELRATLEQTERSRKIAEQELLDASERVOL 672
Qy      543 ---QINPSNDQCKRAEVSFIEFOEKEMEEFVE-----REML 581
Db      673 LHTON---TSLINTKKLETDISOI---QGE MEDIVQEARNAEERAKAKAITDAAMAEEL 726
Qy      582 IKDOEKK--MEDMKGRHHEEIFDLEKEFDEALBOLWYKHG 619
Db      727 KKEQDTSAHLEERMKNLQETVKDLOHRLDBA-BQLALKGG 765

```

Search completed: April 6, 2004, 19:33:25
 Job time : 65 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 19:31:20 ; Search time 21 Seconds
(without alignments)
2862.843 Million cell updates/sec

Title: US-10-030-829-3

Perfect score: 3313

Sequence: 1 MSRRAGPMKEKNVGGYRP.....EPDEALQMLYKKGHLNEDD 625

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :
1: PIR 78:*
2: pir1:*
3: pir2:*
4: pir3:*
5: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	233.5	7.0	644	2	T46211
2	211	6.4	634	2	E96840
3	192	5.8	634	2	E86293
4	186.5	5.6	554	2	E85015
5	183	5.5	559	2	T01724
6	180.5	5.4	895	2	T03446
7	180	5.4	662	2	T01533
8	179.5	5.4	1388	2	S70633
9	178.5	5.4	1957	2	A59294
10	174	5.3	1937	2	I38055
11	173.5	5.2	978	2	A70387
12	173	5.2	488	2	T21596
13	173	5.2	476	2	S00084
14	172.5	5.2	1156	2	B70356
15	172.5	5.2	1407	1	S28589
16	172.5	5.2	1871	2	D96796
17	171.5	5.2	2116	2	A26655
18	171	5.2	1432	2	B85431
19	171	5.2	1948	2	S00485
20	170.5	5.1	725	1	UC5016
21	169	5.1	1938	2	A59293
22	168.5	5.1	1192	2	A71623
23	168.5	5.1	1957	2	A45627
24	167.5	5.1	1898	1	A45973
25	167	5.0	1085	2	F96712
26	166.5	5.0	1979	2	C71622
27	166	5.0	736	2	D86271
28	166	5.0	1088	2	T18559
29	165	5.0	1313	2	A48467

30	165	5.0	1931	2	A59234	slow myosin heavy
31	165	5.0	2663	1	S28261	centromere protein
32	164	5.0	1354	2	S69211	serine/threonine-s
33	164	5.0	1790	2	S67593	transport protein
34	163.5	4.9	980	2	E71606	hypothetical prote
35	163	4.9	1426	2	T00337	hypothetical prote
36	163	4.9	1940	2	A59287	myosin heavy chain
37	162	4.9	1992	1	S02771	myosin heavy chain
38	161.5	4.9	1738	2	T14867	interapkin - slime
39	161.5	4.9	1875	2	S38173	myosin-like protei
40	161	4.9	451	2	G70241	hypothetical prote
41	161	4.9	936	2	S39083	myosin heavy chain
42	161	4.9	1039	2	S18199	myosin heavy chain
43	161	4.9	1837	2	T41023	probable nuclear p
44	160.5	4.8	621	2	S10450	myosin heavy chain
45	160.5	4.8	1642	2	T08880	NMDA receptor-bind

ALIGNMENTS

RESULT 1

T46211
hypothetical protein T8P19.180 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46211
R:Choiene, N.; Robert, C.; Broctier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Sa
Submitted to the Protein Sequence Database, December 1999
A:Reference number: Z231008
A:Accession: T46211
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1644 <CHO>
A:Cross-references: EMBL:AL133315
A:Experimental source: cultivar Columbia; BAC clone T8P19
C:Genetics:
A:Map position: 3
A:Introns: 133/1; 373/3; 403/3; 496/3; 566/3
A:Note: T8P19.180

Query Match	Best Local Similarity	Score	Length
Matches 110; Conservative 88; Mismatches 192; Indels 91; Gaps 19;	7.0%;	233.5;	DB 2;
			Pred. No. 2.8e-05;
QY 181 LASDDYDSVQKSHGSRKQKPFKPFQSLDLSIEQINEPQRQWHCPQNGPQAIW 240	7.0%;	233.5;	DB 2;
7 LSSDDESDISES-----EMDEYGDWYINLKGKLVRLSPQ-AFTCPYCPNKKTSFQ 60	7.0%;	233.5;	DB 2;
QY 241 YNHPILHARTG-----ARRVKLHRELAEVLEKDIQMRGASVTPCGEITYGQWKL 292	7.0%;	233.5;	DB 2;
61 YK--DLLQHASGVGNSNSDKRSASAKASHALATKYVQDDLASASEPSSKQKGNPI 118	7.0%;	233.5;	DB 2;
QY 293 GDEKDYEIWPPMVIIMTRLDKDDNDKWLGMGNBEL-EYFDK-YEALRASHSYGPOG 350	7.0%;	233.5;	DB 2;
119 QDCDHEKLVYPKGIIVNPTTKADGRSAGSSGSKLRREYILRGPNPRVPIWNYLG 178	7.0%;	233.5;	DB 2;
QY 351 HRGMSVLMFESSATGYLAEERLHRELAEWGLDIRIANGQKRSWFSGGVROLYGLATKYDL 410	7.0%;	233.5;	DB 2;
179 HSGTAIVFEMKDWNGHLNGLLPDKAYTVDGHGKDWLKK---DGPGLGYGWIARADY 234	7.0%;	233.5;	DB 2;
QY 411 ---DIFNHSQGRTRLPK-----ELKSVOEMVVKELROISEDNQOLNFKNLSKONKH 461	7.0%;	233.5;	DB 2;
235 NGNNITIGENIRKGTGDLTTAELTEENARKOELVONLRQVLE-----KKKD 281	7.0%;	233.5;	DB 2;
QY 462 AKVLESLEIMSEKLRRTAEDNRIIVRQTMQHEQNREEMDAHREFMDSIKQIHERRDA 521	7.0%;	233.5;	DB 2;
282 MKRIEELCVKSEEL-----NQLMEKKNQOKHYR-ELNATQERTMSHIQIVD--- 330	7.0%;	233.5;	DB 2;
QY 522 KEENFEMLOQOQERAKV-----GQOQONINSSNDCKRAEVSFTIEFOE 568	7.0%;	233.5;	DB 2;
331 DHEKRLTESERKKLEIKCNELAKREVHNGTFRMTL---SBDLONASKNS-DELA 385	7.0%;	233.5;	DB 2;

[illegible]

QY 278 SVIPEGELIYGQMKIGDEKDEKDEIYMPRVLIIMNTRLDKDDNDKXLMGMQGLLETFYDKY 337
 Db 95 EPPRPVPPVPPQOLDDETERPNPHNVYPMWGIYVNPY--KEADDEKLLDSAYMLQTLTKP 152
 QY 338 EALPARHSYCGQHGKSGSVLMFESSATGYLAEARLYHRELAEMGLDIRIAMGQKSMFSG-G 396
 Db 153 KPIEVNAPWVEDSIYGVIAKFNQDMSGFGAGATELEKEFETQSSKKEWTER---SGDS 208
 QY 397 VKQLYGLFATLQD-----DIFNQSQKTRLRKLFELSYGE--MVVVELRLD----SE 443
 Db 209 ESKYVGMFARADDFESQSGPIGYSKGGQLRVLSDISQKNVDORNTVLEELSDMIANTNE 268
 QY 444 DNOQLNYPFNKLS-----KONKHAULVESLEIEMSEKLTRTAEDNNIVHQRITMCH 494
 Db 269 DLNIVQYSYNNLTAMSLQVLRVLDKCKLHLQAFAEYTKKMQGMSLRHI---QKILVDKEKLSN 325
 QY 495 FQNRBEAMDANDRFEMDSIKQIHERRDAKEENFEMLOQDERAKYGGOOONINPSSNDCCR 554
 Db 326 ELDRKMRLESR---AKQL-----EKHEALTELDRQKL-----DEDK 359
 QY 555 KRAEVSFIEFQEKEMEFVEREMLLIKDQKEMDMKKYHNIEIFLEKEPD--EAL 612
 Db 360 RKSDAMNKSLLDASEQKADSVTLVLEHQRQEDALNK---ILLLEKQDTKOTLE 415
 QY 613 -----QLMYKGLHNEDD 625
 Db 416 MEIQELKGLQVM-KH-LGDDDD 436

RESULT 4
 E85015
 hypochetrical protein AT4g01180 [imported] - Arabidopsis thaliana
 C|Species: Arabidopsis thaliana (mouse-ear cress)
 C|Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #next_change 16-Feb-2001
 C|Accession: E85015
 R|Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999

```

388 OKRSMFSG-GVR-OLYGFPLATKOLDIFNOHSOCK-TRKFEELKSJOEWNVKELROISE 443
      ||      |      :|||:|      :|||:|      :|||:|      :|||:|
      :|      :|      :|      :|      :|      :|      :|      :|

```

RESULT 7
T0153
hypothetical protein A_IG005110.22 - Arabidopsis thaliana

A;Experimental source: sex female
C;Superfamily: myosin heavy chain; myosin motor domain homology
F;87-773/Domain: myosin: motor domain homology <MMO>

Query Match	5.4%; Score 178.5; DB 2; Length 1957;
-------------	---------------------------------------

```

OY 157 VEDVDNNAEENBNSDALDDSDDDDLASDYDSDVQKSHGSKONKWFKKPFPGSDLSI 216
Db 1019 VEEDKGNYSNKLKGG--LEQSLDDL-----ENLREKRGGRNIEIKOKRKIGEL--KVAQ 1070
OY 217 EQINEPOROMHCAPCONGPAGAIWMYNLPHLLAHARTKARRV-----KLHR----- 262
Db 1071 ENMEIEIKRQR-----EIE--SNLKKETEAQAITARLEEODVLGKLLKQVNETON 1120
OY 263 ---ELAEVLEKODLQMGASVLPBGSEIYGOMKGLG--DEK-----DYEIWPMPVILMN 311
Db 1121 RITIELEELENEROSRSKAERAKSDLORELEBELGRLDEOGGATAAEOYVNNKREAEALAK 1180
OY 312 TRLD-----KODNDKWLGMNGOELLFVFDKYEALRARRHSYSGPOGHRGS 355
Db 1181 LRRLDEEANNHENOALAIRKKHNDVAVALGQ--TEBOYQKAKATIECKRMOAQORODL 1238
OY 356 VLMFESSATGYLEAEERLHRE--LAEMGLDRIAMQGRSFGSGVROLYGLATK-----QD 409
Db 1239 VAQIDGER-----AERNNNEKLAKQYEMQIA--ELQTKCDBENRROLOEFTTLKRTLSBN 1291
OY 410 LDIFNQSOGK-----TRLKFELKSYQEMVVKELRQISEDNQOLNYFRNKLSKONKH-- 461
Db 1292 TDLGQKQIEEASQVNAWTRLKAQLTSOLEEARRSIDEPEAR-----RNNTAAOMKMYQ 1344
OY 462 -----ANVLEESLEINSEKLRRTAEDNRIVAQ--RTKNQHE--QNRREMDADRFFMOSI 512
Db 1345 HEIEOVLESMEEBELIGSILMKQLSRANAELIQOWOTREFSEBELKGDLEESKRXOMHNM 1404
OY 513 KOIHERBRAKEENFEMLQOOERAKVVGQOOQOININSSNDCKRAAEVSSPIEPOKEWE 572
Db 1405 NELOETLDAANSKISLSLEKTSRLV-----SDDDAQMDEVBRANSVASQVEKKQK 1454
OY 573 EFVEREMLIKDOEKKQEMD-----KKRHHEIIFDLKEKFEDEALBQIMYHGLHN 622
Db 1455 GF-----DKVIDEMWKKKTDIATEVDNAQREARNVSTELFKLSQDEVLLETI--EGJRR 1507
OY 623 ED 624
Db 1508 EN 1509

```

RESULT 10
138055
myosin heavy chain, perinatal skeletal muscle - human
N/Contains: myosin ATPase (EC 3.6.4.1)
C/Species: Homo sapiens (man)
C/Date: 17-May-1996 #sequence revision 17-May-1996 #text change 19-Apr-2002
C/Accession: 138055; JH0154; S12459; S09332; A30220; S49478
R/Jullian, E.H.; Kelly, A.M.; Pompidou, A.J.; Hoffman, R.; Schiaffino, S.; Steedman, H.H.
Eur. J. Biochem. 230, 1001-1006, 1995
A/Title: Characterization of a human perinatal myosin heavy-chain transcript.
A/Reference number: 138055; MUID:95324556; PMID:7601129
A/Accession: 138055
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1937 <RSS>
A/Cross-references: EMBL:Z38133, NID:9558668, PIND:CA86293.1, PID:9558669
R/Karsch-Mizrachi, I.; Peghali, R.; Shows, T.B.; Levinwand, L.A.
Gene 89, 289-294, 1990
A/Title: Generation of a full-length human perinatal myosin heavy-chain-encoding cDNA.
A/Reference number: JH0154; MUID:90323631; PMID:3373371
A/Accession: JH0154
A/Molecule type: mRNA
A/Residues: 1-14, A', 16-859 <KAR>
A/Cross-references: GB:Y00821

A;Experimental source: skeletal muscle
R;Bober, E.
submitted to the EMBL Data Library, January 1989
A;Reference number: S12458

A: Molecule type: mRNA
A: Residues: 502-1071, 'N', 1072-1250, 'DGC', 1253-1376, 'NT', 1379-1913, 'D', 1915-1937 <BOB>
A: Cross-references: EMBL:X51592; NID:g29465; PIDs:CAAS5941.1; PID:g29466
A: Experimental source: clone gMHC-F
A: Biochem. 189, 55-65, 1990
A: Title: Identification of three developmentally controlled isoforms of human myosin heavy chain
A: Reference number: S09332; MUID:90235662; PMID:1691980
A: Accession: S09332
A: Molecule type: mRNA
A: Residues: 502-547, 'X', 549-617, 'X', 619-687, 'X', 689-757, 'X', 759-827, 'X', 829-897, 'X', 899-959, 'X', 961-1031, 'X', 1033-1087, 'X', 1089-1159, 'X', 1161-1231, 'X', 1233-1303, 'X', 1305-1375, 'NT', 1379-1386, 'X', 1388-1456, 'X', 1458-1526, 'X', 1528-1596, 'X', 1598-1666, 'X', 1668-1738, 'X', 1740-1810, 'X', 1812-1882, 'X', 1884-1954, 'X', 1956-2026, 'X', 2028-2098, 'X', 2100-2170, 'X', 2172-2242, 'X', 2244-2314, 'X', 2316-2386, 'X', 2388-2458, 'X', 2460-2530, 'X', 2532-2602, 'X', 2604-2674, 'X', 2676-2746, 'X', 2748-2818, 'X', 2820-2890, 'X', 2892-2962, 'X', 2964-3034, 'X', 3036-3106, 'X', 3108-3178, 'X', 3180-3250, 'X', 3252-3322, 'X', 3324-3394, 'X', 3396-3466, 'X', 3468-3538, 'X', 3540-3610, 'X', 3612-3682, 'X', 3684-3754, 'X', 3756-3826, 'X', 3828-3898, 'X', 3900-3970, 'X', 3972-4042, 'X', 4044-4114, 'X', 4116-4186, 'X', 4188-4258, 'X', 4260-4330, 'X', 4332-4402, 'X', 4404-4474, 'X', 4476-4546, 'X', 4548-4618, 'X', 4620-4690, 'X', 4692-4762, 'X', 4764-4834, 'X', 4836-4906, 'X', 4908-4978, 'X', 4980-5050, 'X', 5052-5122, 'X', 5124-5194, 'X', 5196-5266, 'X', 5268-5338, 'X', 5340-5410, 'X', 5412-5482, 'X', 5484-5554, 'X', 5556-5626, 'X', 5628-5698, 'X', 5700-5770, 'X', 5772-5842, 'X', 5844-5914, 'X', 5916-5986, 'X', 5988-6058, 'X', 6060-6130, 'X', 6132-6202, 'X', 6204-6274, 'X', 6276-6346, 'X', 6348-6418, 'X', 6420-6490, 'X', 6492-6562, 'X', 6564-6634, 'X', 6636-6706, 'X', 6708-6778, 'X', 6780-6850, 'X', 6852-6922, 'X', 6924-6994, 'X', 6996-7066, 'X', 7068-7138, 'X', 7140-7210, 'X', 7212-7282, 'X', 7284-7354, 'X', 7356-7426, 'X', 7428-7498, 'X', 7500-7570, 'X', 7572-7642, 'X', 7644-7714, 'X', 7716-7786, 'X', 7788-7858, 'X', 7860-7930, 'X', 7932-8002, 'X', 8004-8074, 'X', 8076-8146, 'X', 8148-8218, 'X', 8220-8290, 'X', 8292-8362, 'X', 8364-8434, 'X', 8436-8506, 'X', 8508-8578, 'X', 8580-8650, 'X', 8652-8722, 'X', 8724-8794, 'X', 8796-8866, 'X', 8868-8938, 'X', 8940-9010, 'X', 9012-9082, 'X', 9084-9154, 'X', 9156-9226, 'X', 9228-9298, 'X', 9300-9370, 'X', 9372-9442, 'X', 9444-9514, 'X', 9516-9586, 'X', 9588-9658, 'X', 9660-9730, 'X', 9732-9802, 'X', 9804-9874, 'X', 9876-9946, 'X', 9948-10018, 'X', 10020-10090, 'X', 10092-10162, 'X', 10164-10234, 'X', 10236-10306, 'X', 10308-10378, 'X', 10380-10450, 'X', 10452-10522, 'X', 10524-10594, 'X', 10596-10666, 'X', 10668-10738, 'X', 10740-10810, 'X', 10812-10882, 'X', 10884-10954, 'X', 10956-11026, 'X', 11028-11098, 'X', 11100-11170, 'X', 11172-11242, 'X', 11244-11314, 'X', 11316-11386, 'X', 11388-11458, 'X', 11460-11530, 'X', 11532-11602, 'X', 11604-11674, 'X', 11676-11746, 'X', 11748-11818, 'X', 11820-11890, 'X', 11892-11962, 'X', 11964-12034, 'X', 12036-12106, 'X', 12108-12178, 'X', 12180-12250, 'X', 12252-12322, 'X', 12324-12394, 'X', 12396-12466, 'X', 12468-12538, 'X', 12540-12610, 'X', 12612-12682, 'X', 12684-12754, 'X', 12756-12826, 'X', 12828-12898, 'X', 12900-12970, 'X', 12972-13042, 'X', 13044-13114, 'X', 13116-13186, 'X', 13188-13258, 'X', 13260-13330, 'X', 13332-13402, 'X', 13404-13474, 'X', 13476-13546, 'X', 13548-13618, 'X', 13620-13690, 'X', 13692-13762, 'X', 13764-13834, 'X', 13836-13906, 'X', 13908-13978, 'X', 13980-14050, 'X', 14052-14122, 'X', 14124-14194, 'X', 14196-14266, 'X', 14268-14338, 'X', 14340-14410, 'X', 14412-14482, 'X', 14484-14554, 'X', 14556-14626, 'X', 14628-14698, 'X', 14700-14770, 'X', 14772-14842, 'X', 14844-14914, 'X', 14916-15086, 'X', 15088-15158, 'X', 15160-15230, 'X', 15232-15302, 'X', 15304-15374, 'X', 15376-15446, 'X', 15448-15518, 'X', 15520-15590, 'X', 15592-15662, 'X', 15664-15734, 'X', 15736-15806, 'X', 15808-15878, 'X', 15880-15950, 'X', 15952-16022, 'X', 16024-16094, 'X', 16096-16166, 'X', 16168-16238, 'X', 16240-16310, 'X', 16312-16382, 'X', 16384-16454, 'X', 16456-16526, 'X', 16528-16598, 'X', 16600-16670, 'X', 16672-16742, 'X', 16744-16814, 'X', 16816-16886, 'X', 16888-16958, 'X', 16960-17030, 'X', 17032-17102, 'X', 17104-17174, 'X', 17176-17246, 'X', 17248-17318, 'X', 17320-17390, 'X', 17392-17462, 'X', 17464-17534, 'X', 17536-17606, 'X', 17608-17678, 'X', 17680-17750, 'X', 17752-17822, 'X', 17824-17894, 'X', 17896-17966, 'X', 17968-18038, 'X', 18040-18110, 'X', 18112-18182, 'X', 18184-18254, 'X', 18256-18326, 'X', 18328-18398, 'X', 18400-18470, 'X', 18472-18542, 'X', 18544-18614, 'X', 18616-18686, 'X', 18688-18758, 'X', 18760-18830, 'X', 18832-18902, 'X', 18904-18974, '

Db 1610 SRNDALRVKKKMGDINEMEIQLNHNANRLAAESLNRYNTQGIILKETQIHLDDALRGED 1669
 QY 517 -----RRR-----DAK-EENFEMLOOQERAKVVGQOO-----QINPSSND 552
 Db 1670 LKEQALIVERRANLQAEIEELMATEQTERSRKIAEQELDASRVOLHTONTSLINT 1729
 QY 553 CRRAEVSSTFIEFOEKEMEEFVE-----REMLTKDOEKK--MEDM 592
 Db 1730 KKKLNDVSQL-----OSEVEEVIQESRNMEKAKAITDAMMAEELKKEODTSHLRM 1785
 QY 593 KKKHHEIFPLEKEPFDALEQLMYKKG 619
 Db 1786 KKLQOTVKDQHRLEDA-EQLMKKG 1811

RESULT 11
 A70387
 conserved hypothetical protein aq_1006 - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Feb-2001
 C:Accession: A70387
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: A70387
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-978 <ACG>
 A:Cross-references: GB:AE000718; NID:g2983504; PIDN:AC07092.1; PID:g2983515; GB:AE00065
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: aq_1006
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 5.2%; Score 173.5; DB 2; Length 978;
 Best Local Similarity 19.6%; Pred. No. 0.076;
 Matches 10; Conservative 107; Mismatches 212; Indels 215; Gaps 30;

QY 106 NISGRRLSRKYD--NNVAPPPVSRPPLGGMMQARGSAQHTAVQEPFVEDVDN 163
 Db 179 NIEGKEALKKEVELKDYT--PTKGVLE-----KTLKNEELKELKETBEK 225
 QY 164 ASSEENDSDALDDSDILAS-----DDYSDVS--QKSHGRK----- 199
 Db 226 LRQELKAEKESDLRELSQVTKLELNLEKEVEKLEKLEFSRKVAPVPIAKRIEE 285
 QY 200 -----QNKWFKPFSGISLDSI--EQIN-----EPOROMHCPACQNGP 235
 Db 286 IDKLTTELKVRKNTKTELAVLKDSELPQOEBLNRLAEKKEKKEKEKEK----- 337
 QY 236 GAIQWNLHLHARTKGRVYKLRHLEAVLEKDLQWRGASVIPCSEIYQMGKLGED 295
 Db 338 -----LEH-----RLKQLQEIKEIL--KELSLS--LKKK 365
 QY 296 EKQYEIVPWPVIMNTRLDKDNDKVLGMNGOELLEYP-----DKYEALPARSHGPG 343
 Db 366 EREYEAKQEPF-DLSERVEK--GKKLVATEBEKLEKIKELPSEBSYISLAKKELVLE 421
 QY 350 GHRMSVUL-----MFSSATGYLEARLH-----RELAEGLDRIANGQRSMF 393
 Db 422 LQRLKELKEKEGQLENTQTKYKKEKVVHEKVNLEKELERELKEREHLHYHAMVASYS 481
 QY 394 SGGRVQVGLFATKODLIFNHSQGKTRLF--ELKSYDM----- 433
 Db 482 PGDTCFVCGGIYRGKALE--NVDAEGISELHAKELKEKEREIDTTLKYAOKINSLEK 539
 QY 434 -----VKELRQISEN-----QQLNYFQNKLS--KONKHAQVLEE----- 467
 Db 540 EMEKLRNEVEBELKEIPENLKERIKYKLEIKELKELKNTKRYKRLLEBRQKKEBAQA 599

QY 468 -----SLEIMSEKLRRTAEDNRIVRQRTQWQH-----EQNREEMDAHDFPFMSIQIH 516
 Db 600 KLHKAQTELELLKEKIR--EKSRIVKPEKELRYRERLEDVEESLKERINYSKLOIE 656
 QY 517 ERDAAKEENFEMLOOQERAKVVGQOOQONINPSSNDCKRAEVSSTFIEFOE--KEMEEF 574
 Db 657 EKKKIKRKHFEELSSR--KSLLEGELSA--LNEISINLEBERKEKLELANIYVAKSPREV 714
 QY 575 VE-----EREMLIKDOEKEMEDMKKRHH-----EIFPLEKEPFDALEQL 614
 Db 715 VELIYDPKREALERKIKPEESPQSLKKSIEBEKLEYBSQIRLSDIKQEYSEVKTQL 774
 QY 615 MYKH 618
 Db 775 BEKH 778

RESULT 12
 T21596
 hypothetical protein F31d4.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21596
 R:Mortimore, B.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19447
 A:Accession: T21596
 A:Status: preliminary; translated from GB/EMBL/DDAJ
 A:Molecule type: DNA
 A:Residues: 1-488 <WII>
 A:Cross-references: EMBL:Z92832; PIDN:CA807373.1; GSPDB:GN00023; CESP:F31d4.5
 A:Experimental source: clone F31d4
 C:Genetics:
 A:Gene: CESP:F31d4.5
 A:Map position: 5
 A:Intons: 51/3; 193/2; 241/3; 273/2; 459/1

Query Match 5.2%; Score 173; DB 2; Length 488;
 Best Local Similarity 18.1%; Pred. No. 0.035;
 Matches 104; Conservative 100; Mismatches 166; Indels 204; Gaps 23;

QY 124 AP-PPVSRPPLGGMMQARGSAQHTAVQEPFVE-----DDVNASEENDSDALDSD 178
 Db 22 APEPPKNGPPRISPPNTIEKGIILSLSFQ--FSNLTRKKGNDYDLTPEEBGDYG--DEMG 79
 QY 179 DDLASDDYSDVSQKSHGRKQKPFKPFSGSLD-----SLSTEQINEPOROMHCPACQ 232
 Db 80 PVLGVQNY--GILMKYKTKRNRGRARWAKRFVLEKCEFLIYSTSYKVEKTRR----- 131
 QY 233 NGPGALDWNHLHARTKGRVYKLRHLEAVLEKDLQWRGASVIPCSEIYQMGKGL 292
 Db 132 -----IDLHF-----KGI-----IPILGSIYSGDV----- 153
 QY 293 GEDEKYEIVPWPV--IIMNTRLDKDNDKVLGMNGOELLEYPFKYEALPARSHGPG 350
 Db 154 --DKKNGLLAHQPLSAIIVASDHQTOEMML-----KALNS----- 189
 QY 351 HGRMSVLMFSSATGYLEARLHRELAEMGL-----DRIAWGQRSMFSGGVROL 400
 Db 190 ---ATKISYKNTAVG--ETMIRELNGVILNEEKSYEERLEAEK----- 231
 QY 401 YGFLATKODLIFNHSQGKTRLF-----LKSIOEMVVELQISEDNQQLNY 450
 Db 232 ---ARKEHBRADDELAKKELEAEBEKLIPTTKLKDQLVNKLKMTNEMKTKTLBO 287
 QY 451 FKNTLSKONKHAQVLESLEIMSEKLRRTAEDNRIVRQRTM----- 492
 Db 288 EKMSINSTEHLQNMESLINIEKELHEQLQF--IVAREKVLIDNQNISTDCQQLNRL 345
 QY 493 -----OHEQNRREEMDAHDF-----MDSIKQIHERDAE 523
 Db 346 MEIETSRNCINTREKELINLKNMEQKQDLEKERYQYTKTSELMDHLKEVSEBQDLE 405

QY 524 ENEEMLQOQERARVGGQOQINPSSNDCKRAAEVSSFFIEQEKEMEEFVEREMLIK 563
 DB 406 SE---LKEQMMARLGAERKQ-----LQAEALREHL-----EWALK 437
 QY 584 DQEKMEDEMKKRHHHEIFDLKEKFEDEALQOLMYK 617
 DB 438 MTGAQMTELOEHIMPVHKLREFFEQCAESRPE 471

RESULT 13

S00084
 myosin heavy chain, fast skeletal muscle - rabbit (fragment)
 C/Species: Oryctolagus cuniculus (domestic rabbit)
 C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
 C/Accession: S00084; GI:84807
 R/Maeda, K.; Sczakiel, G.; Wittlinghofer, A.
 Eur. J. Biochem. 167, 97-102, 1987
 A/Title: Characterization of cDNA coding for the complete light meromyosin portion of a
 A/Reference number: S00084; PMID:87304245; PMID:3305014
 A/Accession: S00084
 A/Molecule type: mRNA
 A/Residues: 1-676 <MAE>
 A/Cross-references: EMBL:X05958; NID:g1622; PIDN:CAA29391.1; PID:g1364242
 A/Note: the sequence from Fig. 5 is inconsistent with that from Fig. 3 in having 561-Arg
 R/Maeda, K.; Roessch, A.; Maeda, Y.; Kalbitzer, H.R.; Wittlinghofer, A.
 FEBS Lett. 281, 23-26, 1991
 A/Title: Rabbit skeletal muscle myosin. Unfolded carboxyl-terminus and its role in molec
 A/Reference number: S14807; PMID:91200294; PMID:2015900
 A/Accession: S14807
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 668-676 <MA2>
 C/Superfamily: myosin heavy chain; myosin motor domain homology
 C/Keywords: actin binding; ATP; coiled coil; muscle; skeletal muscle
 F/21-676/Domain: light meromyosin <LME>

Query Match 5.2%; Score 173; DB 2; Length 676;
 Best Local Similarity 21.1%; Pred. No. 0.052;

Matches 124; Conservative 91; Mismatches 184; Indels 188; Gaps 30;

QY 141 ARGGSAGHTAVQEPFVEDVDVNASBEEND---SDALDSS---DDLASDDYDSVSQK 193
 DB 43 SRGSG-----QAFITQOIEGLKRLQLEETKAKSALAHALQSSRRDCLLEQVEEBOEAK 95
 QY 194 SHGRKONKPKFKFPSLSLSTEQINEPQRMHCAPACONGPAILMYNHLPLAHARTK 253
 DB 96 AEIQRAMSK-----ANSEVSQWR-TYCEET-----DAIQTEELBEKCKK 133
 QY 254 GARRVHLRELAVLEKDLQMRGASVIPCGEIYGQWKGJGDEKDYEIWPPMVIIMNT- 312
 DB 134 LMQRLDQAEHVAANS-----KCAELEKTKQRLQNEADL-----MIDVERSN 177
 QY 313 ---RLDKDNDKWLGMGNQELLEFDKYEALRAHSYGPQGRGMSVLMFE-----SSA 363
 DB 178 ATCARMDKKQRN-----FDKVLAEWKIKYEETQALEASQKESRSISTEVFKVKNAYEES 232
 QY 364 TGYLFA---EHLHRELAMGJDRITAMG-----QKRSMSGGVRQLYGF 493
 DB 233 LDHLETLTKENKVLQOEIISL-TEQIAESAKHILHEKYKKQIQDQESSELQALAELEGSS 291
 QY 404 LATKQDLDIFNQSQCK-TFLKELFSYQEMVVKELQISDNQOLNYFKNLKSQNKHA 462
 DB 292 L-----EHEBCKILRIQLELNOVSEIDRKIAKDEBEIDQL-----KEN-HL 332
 QY 463 KYLEESLEIWMSEKLRATBEDNRIVRQTKQHEQNEENDAH--DRPFMDSIK----- 513
 DB 333 RVEVSMQSTLDAEIR---SRNDALRIKKKMEGDLNEMEIQLNHNANROAIAIKLNTQOG 389
 QY 514 ---QIH-----ERR---DAK-EENFEMLQOQERAKVVGQO--- 542
 DB 390 ILKDTQLHLDVAGDDHKEQALAVERRKNMLQALIEELIRASLEQTERSRVADDDLLD 449
 QY 543 ---QINPSSNDCKRAAEVSSFFIEQEKEMEEFVEE----- 577

DB 450 ASERVQLLHTON---TSLINTKKKLETTDISQI---QGMEDIVQEARNAEAKKAITDA 503
 QY 578 ---REMLIKDOEKK--MEDMKKRHHHEIFDLKEKFEDEALQOLMYKHG 619
 DB 504 AMMAEELKKEODTSAHLEBKRMQMEQTVKQLOOLDEA-EQALAKGG 549

RESULT 14

B70356
 chromosome assembly protein homolog - Aquifex aeolicus
 C/Species: Aquifex aeolicus
 C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Jun-2000
 C/Accession: B70356
 R/Beckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
 V.
 Nature 392, 353-358, 1998
 A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A/Reference number: A70300; PMID:98196666; PMID:9537320
 A/Accession: B70356
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-1156 <AGP>
 A/Cross-references: GB:AE000699; NID:g2983238; PIDN:AAC06839.1; PID:g2983243; GB:AE00065;
 A/Experimental source: strain VFS
 C/Genetics:
 A/Genes: xcpC
 C/Superfamily: chromosome segregation protein SMCI

Query Match 5.2%; Score 172.5; DB 2; Length 1156;
 Best Local Similarity 19.9%; Pred. No. 0.11;

Matches 82; Conservative 76; Mismatches 139; Indels 115; Gaps 15;

QY 268 LEKDLQMRGASVIPCSEI-----YQGMKGL-----GDEKDYEIWPPMVI 308
 DB 555 LAKRMNLGRFSFPLNVRVVERPRLYPTKGAVDPAVNLVEYDPKFEKVKVKGFGDTLI 614
 QY 309 IMNTRLDKDNKWLGMGNQELL---EYFDKYEALRAHSYGPQGRGMSVLMFESSAT 364
 DB 615 VENN-----ESAKAIDIGIRNVTTLGELFEKSGV-----TGAV-----KSS 653
 QY 365 GYLAERLHRELAEWGLDIRIANGQKSMFSGGVRQLYGLATKQDLDIFNQSQKTRIK 424
 DB 654 GELNKKRYEEELQRLNAEBEKLKNEBSIIQKIRIRIHLISEKTAL-----LK 701
 QY 425 FELKSYQEMVVKELQISDNQOLNYFKNLKSQNKHAVALLESLEIMSEKLRATBED-- 482
 DB 702 VSEBKIEELSSGLEQYERK-----FKKLENSKEYLILBEKCLINVEDKLEALEEIE 755
 QY 483 ---NRIVRQTKQHEQNEENDAHDRPFMDSIKQIHE-----RDKAKENPFM 528
 DB 756 YIEBKLNINLKLKGGDIKRRYSRGEVBEKRRYSKVRKQVSEIKSLNLEIRLNTKTYEL 815
 QY 529 ---LQOQERAKVVGQOQOINPSSNDCKRAAEVSSFFIEQEKEMEEFVEREMLI 582
 DB 816 EYLEKEILOEKER-----EYLTERISLKKKEIENLILFKKTLQE-VKEAEVAV 864
 QY 583 KQOEKMEDEMKK-----RHHEIFDLKE---FDALQOL 614
 DB 865 YVYIKQKELEKEIINLAKSKLKLKIKYEBELKEKFEKEKNLKVLEKIENTL 916

RESULT 15

S28589
 trichohyalin - rabbit
 C/Species: Oryctolagus cuniculus (domestic rabbit)
 C/Date: 12-Mar-1993 #sequence_revision 01-Mar-1996 #text_change 22-Jun-1999
 C/Accession: S28589
 R/Pietz, M.J.; Rogers, G.E.
 submitted to the EMBL Data Library, December 1992
 A/Description: Examination of the gene encoding rabbit trichohyalin.
 A/Reference number: S28589
 A/Accession: S28589

A: Molecule type: DNA
A: Residues: 1-1407 <EF2>
A: Cross-references: EMBL:219092; NID:91746; PID:CAA79519.1; PID:91747
C: Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root sheath. Covalent modifications to this protein include conversion of arginine to citrulline and
C: Genetics:
A: Introns: 46/3
C: Superfamily: trichohyalin; calmodulin repeat homology
C: Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
F: 49-81/Domain: calmodulin repeat homology <EF2>

Query Match 5.2%; Score 172.5; DB 1; Length 1407;
Best Local Similarity 21.08; Pred. No. 0.13;
Matches 87; Conservative 71; Mismatches 116; Indels 141; Gaps 17;

```
QY 278 SVIPGGEIYGOM-----KGIGDEKDYELVW-----PPMVIIMNTRLDKDN 319
DB 7 SIIDITIFPNQYASHDCGAVLKKKDKLIDREFGAVLQRPHPETVDVMLELDR-DS 65
QY 320 DKWLGMGNQELLEYFDKYEALRARSYGPQGHGMSVLMPESSATGYLAEERLHRELAE 379
DB 66 DGLVG-----FDEP-----CLIFKLAQAAYVALGQ-----AS 93
QY 380 GLDRIAMGQKRSFSGVRLYGLATKODLIDFNHSGKTRL-----KTELKS 429
DB 94 GLDE-----EKRS-----HGEKGRLQNRQEDQRRFELRD 125
QY 430 YQ-----EMVVKELRQISEDNQQLNFKNLASKO--NKHAKVLEESLE----- 470
DB 126 RQFDEPERRRWQOEERELAEERERKRRERFQHSRYRDKERQRLQRELEERRAE 185
QY 471 -----IMSEKLRRTAEDNRIVORTKMQHEQNREEMDAHDHFFMDSIKQI 515
DB 186 EQQLRRRKRDSEEFIEERQLR--REQQELKRELREERQORRRERERQHERALQEEERQL 243
QY 516 HERRDAKEENFEMLQOEERAKVVGQOQINPSSNDCKRAEYVSFIEFOE-----KEM 571
DB 244 LQQRWRREPRQOQLRRELIEIRERQRLQEERERQOQLRQRLQEERERQOQLRREL 303
QY 572 EEVFERERMLIK-----DOEKKMEDMKRRHHEIFDLEK--EPDEALEQLM 615
DB 304 EE-IRERERQRLQEERERQRLQEERERQOQLKRELIEIRERQRLQEERERQOQL 357
```

Search completed: April 6, 2004, 19:35:24
Job time : 24 secs